BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : A61K X

A2

(11) International Publication Number: WO 90/07321

(43) International Publication Date: 12 July 1990 (12.07.90)

(21) International Application Number: PCT/US89/05067

(22) International Filing Date: 10 November 1989 (10.11.89)

(30) Priority data:

289,201	23 December 1988 (23.12.88)	US
315,736	24 February 1989 (24.02.89)	US

(71) Applicant: THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY [US/US]; Stanford University, Stanford, CA 94305 (US).

(72) Inventors: WEISSMAN, Irving, L. ; 1170 Welch Road, Apt. 711, Stanford, CA 94305 (US). HOLZMANN, Bernard ; Gufidaunerstr. 30, D-8000 Münich 90 (DE). SIEGELMAN, Mark, H. ; 2035 Columbia Avenue, Palo Alto, CA 94306 (US).

(74) Agent: LEYDIG, VOIT & MAYER; 350 Cambridge, Suite 200, Palo Alto, CA 94306 (US).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: HOMING SEQUENCES AND THEIR USES

(57) Abstract

Proteins are identified as homing receptors for Peyer's patches and lymph nodes, where the proteins may be used for inhibiting homing of lymphocytes or providing for homing of drugs or other compositions for *in vivo* diagnosis or therapy. In addition, nucleic acid compositions are provided which may be used for expression of the proteins or fragments thereof or for transforming cells to provide for enhanced homing capability or for inhibiting or modulating such homing.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MR	Mauritania
BE	Belgium	GA	Gabon	MW	Malawi
BF	Burkina Fasso	GB	United Kingdom	NL	Netherlands
BG	Bulgaria	HU	Hungary	NO	Norway
BJ	Benin	IT	Italy	RO	Romania
BR	Brazil	JP	Japan	SD	Sudan
CA	Canada	KP	Democratic People's Republic of Korea	SE	Sweden
CF	Central African Republic			SN	Senegal
CG	Congo	KR	Republic of Korea	SU	Soviet Union
CH	Switzerland	LI	Liechtenstein	TD	Chad
CM	Cameroon	LK	Sri Lanka	TG	Togo
DE	Germany, Federal Republic of	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

HOMING SEQUENCES AND THEIR USES

INTRODUCTION

Technical Field

The technical field of the subject invention concerns physiologically active proteins associated with cellular homing to target anatomical sites.

Background

The immune system, unlike most organ systems which are consolidated in one anatomical location, is dispersed over an entire organism. It exists as circulating elements in the blood, through which it gains access to nearly all body tissues, and as innumerable lymphoid aggregates throughout the body. Therefore, the immune system is placed under a special constraint, which is managed by substituting extensive cell-cell recognition and interactive events.

The constraints imposed by a physically unmoored blood-borne immune system containing a particular antigen reactive lymphocyte at very low frequency demands additional organization to insure appropriate interaction with antigen regardless of the antigens portal of entry. The dynamism of the circulating lymphoid system is relieved by scattered solid collections of lymphoid elements, such as thymus, lymph nodes, Peyer's patches, and spleen, which together constitute the lymphoid organs.

Perpetual percolation of lymphocytes through lymphoid organs efficiently arms each of these organs with the entire repertoire of antigen-reactive cells; lymphocytes recirculate from blood to lymphoid organs and back to blood, generally passing the efferent

lymphatic vessels and their collecting ducts. The specific portal of entry of lymphocytes from bloodstream into peripheral lymphoid organ was identified as specialized postcapillary venules bearing unusally high-walled endothelia, subsequently designated high endothelial venules (HEV's). Recirculating lymphocytes, but not other blood-borne cells, specifically recognize, adhere to luminal walls, and migrate through this highly specialized endothelium into the lymphoid organ parenchyma proper. This migration of recirculating lymphocytes from blood stream to particular lymphoid sites has been called "homing," and the cell surface structures mediating recognition and adherence to lymphoid organ HEV's have been called "homing receptors." Therefore, lymphocyte homing appears to be regulated by the expression of complementary adhesion molecules on each of the two participants, the recirculating lymphocyte and the specialized lymphoid organ HEV's.

The homing phenomenon is an important aspect of many systems, both for the benefit and detriment of the host. The ability to home specific cells to particular organs can be of great benefit in the defense of disease, particularly where the cells may be introduced adjacent to the particular organ of interest, so that the specialized cells will populate that organ. By contrast, in the case of cancer, particularly lymphomas, the homing receptor may serve to enhance metastases, so as to spread the neoplasia throughout the immune system. Homing may be an aspect of the inflammatory response, which may result in autoimmune diseases. The ability to diminish the inflammatory response or attack on native tissue may serve as a therapy in the case of such diseases as rheumatoid arthritis. There is, therefore, great interest in being able to identify the molecules involved with homing, the mechanisms by which homing

occurs, and means for modulating the homing response.

Relevant Literature

Reviews of the integrin family of proteins may be found in Hynes (1987) Cell 48:549-544; and Ruoslahti and Pierschbacher (1987) Science 238:491-497. Descriptions of the VLA family of proteins are provided by Takada et al. (1987) Proc. Natl. Acad. Sci. USA 84:3239-3243; Hemler et al. (1987) J. Biol. Chem. 262:3300-3309; Hemler et al. (1987) J. Biol. Chem. 262:11478-11485; and Hemler et al. (1988) Immunol. Today 9:109-113.

The structure of the alpha and beta subunits has been described by Kishimoto et al. (1987) Cell 48:681-690; Argraves et al. (1987) J. Cell Biol. 105:1183-1190; Fitzgerald (1987) J. Biol. Chem. 262:3936-3939; Suzuki et al. (1986) Pro. Natl. Acad. Sci. USA 83:8614-8618; Poncz et al. (1987) J. Biol. Chem. 262:8476-8482; Arnaout et al. (1988) J. Cell Biol. 106:2153-2158; Pytela (1988) EMBO J. 7:1371-1378; Corbi et al. (1987) EMBO. J. 6:423-4028; and Corbi et al. (1988) J. Biol. Chem. 263:12403-12411.

A description of the MEL-14 antibody and the lymph node specific homing receptor to which it binds is described by Gallatin et al (1983) Nature 304:30-34; Seigelman et al. (1986) Science 231:823-829; St. John et al. (1986) Science 231:845-850; Jalknen et al. (1986) Eur. J. Immunol. 16:1195-1202; and Jalknen et al. (1987) J. Cell Biol. 105:983-990.

See also Dailey et al. (1982) Proc. Natl. Acad. Sci. USA 79:5384, which suggests that CTL's specific for a particular cell which does not have a homing receptor must be in the drainage of the target for activity.

SUMMARY OF THE INVENTION

Methods for modulating homing to peripheral lymphoid organs, e.g., lymph nodes and mucosal lymphoid organs, e.g., Peyer's patches, are provided, employing antibodies to the homing receptor core proteins, nucleic acid compositions for the expression of core proteins, methods of transfecting cells to provide homing capability, and the use of the various compositions in diagnosis and therapy. Particularly, mouse and human alpha and beta subunits of the integrin family used for homing to mucosal lymphoid organs and lymph node homing receptors are described.

DESCRIPTION OF THE SPECIFIC EMBODIMENTS

Methods and compositions are provided which find use in the modulation of homing of cells to peripheral lymphoid organs, e.g., lymph nodes and mucosal lymphoid organ and/or membrane sites, e.g., Peyer's patches in mammalian hosts, particularly human hosts. It is now shown that VLA-4 is a member of the integrin family associated with homing to the high endothelial venules (HEV's) associated with Peyer's patches, while a ubiquinated protein which is highly glycosylated is shown to be associated with a lymph node homing receptor.

In accordance with the subject invention, nucleic acids encoding the core proteins or physiologically active fragments thereof, the use of such nucleic sequences for transfection of cells to provide homing to the particular sites or produce peptides which may be used as antagonists, the proteins and fragments thereof which may be used as antagonists, antibodies to the proteins, and antidiotypes are described.

The various compositions may be used in a variety of ways: In diagnosis, to define the presence or absence of cells, tissue or bodily fluids containing

and/or expressing the homing receptors or the complementary ligand; in therapy, to enhance the homing phenomenon by enhancing the population of homing receptor on the surface or to inhibit the homing phenomenon by either employing competitive proteins as antagonists, or antibodies which may inhibit complex formation between the homing receptor and its complementary ligand, particularly in relation to an imflammatory response; in research, to identify the HEV proteins binding to the different domains of the homing receptors and the effect of mutations in the domains on binding.

In addition, the subject peptide compositions may be used to direct various compositions to particular sites in a mammalian host, by binding the moiety of interest to a subject peptide. In this way, greater specificity of association between the moiety of interest and the high endothelial venule or other cellular targets may be achieved.

The nucleic acid sequences may be used for producing the subject peptides, or fragments thereof, in accordance with genetic techniques or may be joined to other nucleic acid sequences, under conditions involving a replicating species, where the conditions provide for expression of the subject peptides jointly with other proteins, thus directing the replicating species to the target sites.

First will be considered the mucosal lymphoid tissue and organs, including Peyer's patch, homing receptors, associated with the integrin family, where both the mouse and human proteins will be described. It is understood that the mouse and human proteins find analogy, one with the other, in that these proteins are immunologically cross reactive, and that there is substantial conservation of these sequences in the two species. However, the proteins of the two species are given different names and until a common nomenclature

is provided, the different names and their analogues will be considered.

The mouse proteins which are described are referred to as LPAM-1 and 2, where LPAM stands for lymphocyte Peyer's patch HEV adhesion molecule, while VLA stands for Very Late Antigens (of lymphocytes). LPAM-1 and -2 share a common alpha unit referred to as α_{4m}, which binds to two different beta subunits, where the beta subunit of LPAM-1 is referred to as β_p which does not find analogy with the heretofore reported beta subunits of the human integrin VLA proteins, β_1, β_2, and β_3. The LPAM-2 beta unit finds analogy with integrin β_1. The LPAM proteins are further characterized in that LPAM-1 is a heterodimer of α and β subunits of about 160kd and 130kd M_r respectively, that the association requires the presence of calcium ions, and that proteins of 84kd M_r and 62kd M_r present in LPAM-1 precipitates appear to be products of proteolytic immunoprecipitates of proteolytic processing of alpha chains. The structure of LPAM-1 is virtually identical of that of the human integrin receptor VLA-4, with cross-reactivity of monospecific antisera between the alpha units of the VLA-4 and LPAM-1 proteins being observed.

The LPAM-1 and -2 proteins and their subunits are provided in purified form, generally being at least 50 wt.%, usually at least about 90 wt.% preferably at least about 99 wt.%, particularly as to the presence of other proteins. The compositions may be present in lyophilized form, in solution, or formulated with other components, as desired.

The alpha and beta subunits are transmembrane glycoproteins with large extracellular and short cytoplasmic domains. The human beta subunits show 40-48% identity to each other. In amino acid sequence, their extracellular domains contain 56 cysteine residues, all of which are conserved. The alpha

subunits of integrins contain a series of sites capable of divalent cation binding, show substantial amino acid sequence similarities between the various alpha subunits and in some instances consist of two disulfide linked polypeptides. The cysteines of the beta subunits include 4 repeats of an 8 cysteine motif. Rather than the heretofore observed combination of subunits involving a single beta subunit binding to a number of different alpha subunits to provide different homing receptor molecules, it has now been discovered that a single alpha subunit may bind to different beta subunits to provide for different homing receptors, having overlapping homing capability. Thus, individual alpha subunits may be combined with different beta subunits to produce homing receptors having overlapping, but different binding profiles.

The lymph node homing receptor binds to the antibody MEL-14 which may also recognize a ubiquitin epitope. The lymph node receptor is characterized by being a highly glycosylated protein which is also ubitiquinated and has a core structure as described in the experimental section. The precursor protein has an unusually long signal sequence, which has the normal hydrophobic region, which in turn is followed by a hydrophilic domain. The molecular weight of the glycosylated protein is about 90 kD, while the ubiquitin-free core protein is about 35-40 kD. The mature protein has a pI of about 4-4.5 (See Siegelman and Weissman, Ubiquitin, ed. Martin Rechsteiner, Plenum Publishing Corp., 1988, chapter 9, pp. 239-69).

Murine and human lymph node homing receptors have the nucleic acid coding and flanking sequences and related amino acid sequence as described in the Experimental section.

The murine cDNA clone described in the Experimental section has a 54 bp 5' untranslated region followed by an initiator ATG codon, which begins an

uninterrupted open reading frame of 1,116 bp. The reading frame encodes a protein with a hydrophobic leader sequence 38 amino acids in length, before reaching the initial tryptophan residue of the mature protein. The leader sequence has a length unusual for a signal sequence.

The mature protein possess 10 potential asparagine-linked glycosylation sites consistent with protein characterization studies which show extensive glycosylation in endoglycosidase F digestion. These are contained within an identical repeat unit structure. The mature protein contains 22 cysteine residues, where 12 of the cysteines are present in a complement regulatory repeat structure and an additional 9 cysteines are concentrated in the 60 amino acids just preceding the repeat units involving the EGF-like domain. This results in a highly cysteine-rich pre-transmembrane region of 180-190 amino acids.

The deduced mature protein is 334 amino acids in length with a calculated molecular weight of 37,600. The hydrophobic transmembrane regions encompassing amino acids from about 295-317 is followed by a cluster of positively charged residues and a hydrophilic cytoplasmic tail of 18 amino acids. A hydropathy plot further shows distinct regions of relative hydrohilicity, concentrated in the amino-terminal 150 acids and in the membrane proximal approximate 20 amino acids. The intervening extracytoplasmic portion is comprised of a relatively electrically neutral stretch which is characterized by repeat units, identical at nucleotide as well as protein level.

The extracytoplasmic portion of the receptor is made of three separate extracytoplasmic domains, defined by their homology to three disparate protein motifs. One shows homology to the carbohydrate binding domains of animal lectins (positions 74-118); the succeeding 37 amino acids (positions 119-155) occupy

the region between the lectin domain and complement regulatory repeat units and exhibit similarity to the epidermal growth factor (EGF) cysteine-rich repeat unit; the third region is comprised of two identical repeat units conforming to the consensus sequence of homologous repeats found in complement regulatory and other proteins (positions 156-217).

The individual domains may serve for their respective purposes as separate and distinct entities. For example, the lectin domain may be used for binding to a complementary sugar or identifying sugars with the particular domain. The EGF domain may be used to bind to the EGF receptor, competing with natural EGF for binding to the receptor. The complement regulatory repeat units may be used in regulating complement, by being combined with the members of the complement cascade to modulate complement formation and lysis.

The EGF-like domain preserves many of the cys-gly residues characteristic of the EGF repeat unit, with six consensus cysteines present, as well as glycines at 147 and 150, and tyrosine at 148. The relationship of these conserved residues is identical to that of human and bovine blood clotting factors IX and X, and the Drosophila Notch gene product, and similar to other molecules containing EGF-like domains. These regions are believed to be involved in cell-cell interaction mechanisms essential for embryonic differentiation of ectoderm into neural and epidermal precursors. The EGF-like domain further shares homology with a portion of one of the cysteine-rich repeat units of the β-chain of the integrin LFA-1 β_2-chain in the human.

The duplicated repeat unit has 62 amino acids in length and spans positions 156-217 and 218-279. A known protein exhibiting significant homology to this sequence is the murine complement factor H, a serum protein with complement regulatory activity. The same

homologous repeat motif exists in a number of complement regulatory proteins which bind C3/C4, and in other proteins such as Il-2 receptor, the β_2-glycoprotein serum protein and factor XIII.

The lymph node homing receptor will be substantially conserved among the various mammalian species. Thus, the receptor will have a signal sequence, lectin-like domain, and EGF-like domain, and a repeat sequence, where the repeat finds homology with complement regulatory proteins. These various domains may serve to provide for individual activity, being agonists or antagonists as to their particular functions. The sequences may be used to inhibit binding of the homing receptor to the HEV.

The sequences may be modified, where a sequence of only about 8 amino acids may be employed coming within one of the sequences of the various domains. The sequences may be mutated, by changing up to 20% of the amino acids, more usually not more than about 10%, where deletions and insertions may involve from about 1 to 10, usually from about 1 to 5 amino acids.

The DNA sequences corresponding to the various domains may be used as probes for finding other proteins having like domains, sharing homology in function with the domains of the homing receptor.

Depending upon the particular protein employed, different sites for homing may be achieved. In the case of the LPAM or VLA proteins described above, homing will be primarily to mucosal tissue, which includes Peyer's patches, appendix, tonsils, adenoids, bronchial mucosa, mesenteric lymph nodes, or the like. For the peripheral lymphoid organ homing receptor, all peripheral lymph nodes, and potentionally the spleen, will be the primary targets.

The subject proteins, nucleic acid sequences encoding the proteins, or chemically, biologically or

physiologically active or useful fragments thereof may find a variety of applications. The proteins or fragments thereof may be used to produce antisera or monoclonal antibodies specific for one or more epitopes of the subject proteins. In turn, the antibodies may be used to produce anti-idiotype antibodies which may directly compete with the homing receptor for binding to the complementary ligand. These antibodies find use in inhibiting the complex formation between the homing receptor and its complementary ligand. Thus, the antibodies may be used to prevent homing of cells to mucosal sites or lymph nodes. The inhibition of homing may find use in the treatment of inflammatory bowel diseases such as regional ileitis, ulcerative colitis, severe lymphadenitides, histiocytic disorders of lymph nodes or other inflammatory conditions. The antibodies may be used to inhibit metastases, where a neoplastic condition is associated with transport to mucosal sites or lymph nodes.

The proteins or fragments thereof, capable of binding to the complementary ligand may also be used as antagonists for complex formation. Thus, by administering the homing receptor protein or fragment thereof to a host, the protein may serve to home to the complementary ligand and inhibit the binding of the homing receptor associated with the target cells.

Rather than acting as inhibitors to prevent complex formation between lymphocytes and HEV's, the proteins, fragments thereof, or anti-idiotype antibodies may serve to direct a wide variety of molecules to the homing site. Thus, in the case of neoplastic tissue, by administering one of the subject compounds or compositions bound to a therapeutic drug, one can direct the binding of the therapeutic drug to the desired site for retention and concentration at the desired site. One could provide for the binding of radioisotopes for <u>in vivo</u> diagnosis or imaging, for

radiotherapy, or the like. Alternatively, one could bind cytotoxic drugs, either directly or in the lumen of liposomes, where the subject protein would direct the cytotoxic drug to the homing site.

The nucleic acid sequences encoding the proteins of the subject invention will usually be at least 12nt, more usually at least 16nt, and may be 50nt or more, providing for a sequence different from the members of the homing receptor proteins having substantially different target profiles from the same or different species. The DNA sequences will be present as other than a mammalian chromosome, generally present as less than 50knt, particularly during manipulations, such as cloning and constructions. If introduced in a cell, the sequence may be integrated in the chromosome, but may be at other than its natural site in the genome. The sequence may be a genomic sequence, comprising all or part of the structural gene or a cDNA comprising all or part of the coding sequence.

The sequences may be identical to the sequence of the gene or be different, including transitions, transversions, deletions or insertions. For use in detecting sequences encoding proteins having analogous function, related sequences may have as little as 30% homology, usually at least about 40% homology. For mutant sequences or closely related proteins, there will usually be at least about 95% identity with the wild-type sequence, particularly conservative substitutions, although there may be substitutions which result in fewer than 5% changes in amino acids, usually not more than a total of 10 amino acids, preferably not more than about 5 amino acids.

The nucleic acid sequence may be modified by being labeled with a label capable of providing a detectable signal, either directly or indirectly, such as a radioisotope, biotin, fluorescer, etc.

The nucleic acid sequences encoding the subject proteins or fragments thereof may be used for expression of the peptides. Thus, vectors may be prepared which provide for expression of a peptide of interest, which may then be harvested for use as described above. A large number of expression vectors are commercially available or have been described in the literature for expression in a variety of prokaryotic and eukaryotic hosts. Hosts of interest include E. coli, B. subtilis, yeast, such as Saccharomyces, Kluyveromyces, etc., filamentous fungi, such as Neurospora, mammalian cells, such as CHO, COS, HeLa cells, L cells, immortalized T- or B-cells, e.g., EBV immortalized B-cells, etc. Replication systems include ColE1, simian virus 40, baculovirus, lambda, 2mμ plasmid, bovine papilloma virus, etc. A large number of transcription initiation and termination regulatory regions have been isolated and shown to be effective in the transcription and translation of heterologous proteins in the various hosts. The literature is replete with examples of these regions, methods for isolating them, and their manner of manipulation, and such disclosure will not be repeated here.

Vectors may be prepared which will usually include one or more replication systems for cloning or expression, one or more markers for selection in the cellular host, e.g., antibiotic resistance, and one or more expression cassettes for expression of the subject proteins. Desirably when expressing the subject proteins in a cell to be used for homing to a target site, regions other than the wild-type transcription initiation region will be used, where the initiation may be constitutive or inducible, but not subject to the wild-type regulation.

The coding sequences may be synthesized, isolated from natural sources, may be prepared as

hybrids, or the like. Joining of the coding sequences to the transcriptional regulatory sequences may be achieved by restriction, ligation, use of adaptors or polylinkers, or the like. The particular method of preparing the expression vector, introducing the vector into an appropriate host, growing the host, whereby the subject peptide is expressed, and then isolating the subject peptide is not critical to this invention and any convenient technique or protocol may be employed.

Besides introducing an expression cassette comprising the subject coding sequences for producing the protein, in many situations it will be advantageous to transform cells to enhance their capability for homing or impart to a cell a homing capability. Thus, it may be of interest to transform stem cells, usually syngeneic or allogeneic, and cultivate the stem cells to produce stem cells of a particular lineage or subset, such as natural killer cells, tumor infiltrating lymphocytes, cytotoxic T-lymphocytes, B-cells, or the like. One could then provide for the homing of these cells to a particular site or sites, where these cells provide a desired function. Alternatively, one could isolate precursor cells, e.g., $CD4^-$, $CD8^-$, or mature cells, e.g., $CD4^+$ or $CD8^+$, and transform them in an analogous manner. The cells could then be returned to the host for appropriate therapy.

Depending on the choice of host, one could obtain the core protein (unubiquinated) or the ubiquinated protein. Using microorganism hosts or other eukaryotic hosts which do not have the processing capability to ubiquinate the core protein will result in a product which is unprocessed. By contrast, by using an appropriate host, the ubiquinated product will be obtained.

The signal sequence of the lymph node homing receptor may also be used for transport of a wide variety of proteins along particular pathways of

intracellular trafficking to result in special post-translational modifications for placement in various intracellular compartments or into the nutrient medium. Thus, the subject signal sequence provides an additional signal sequence which may find preferred application with certain proteins.

The a_{4m} or b_p protein may be used to obtain the gene encoding the a_{4m} or b_p protein, either as the genomic gene or as cDNA. By preparing a probe based on an amino acid sequence of the a_{4m} or b_p protein of at least about 6, preferably 8, amino acids, using the redundancy of the codons to prepare all possible variations, one can identify sequences in a library comprising either cDNA or genomic DNA. The cDNA library may be prepared in accordance with conventional ways from cytoplasmic RNA from a homing Peyer's patch HEV binding lymphoma, e.g., TK1, and then subtracted with a T-cell lymphoma which does not home to Peyer's patches. The subtracted library may then be probed with the probe indicated above. Positive clones may then be sequenced to identify the presence of a nucleic acid sequence encoding the correct amino acid sequence. If necessary, where a complete structural gene sequence is not obtained, the truncated sequence may be used as a probe to identify a clone having a complete sequence or, if necessary, to use the truncated sequence as a primer for reverse transcription of mRNA from the original source. Once a complete sequence has been identified, the sequence may then be used in a variety of ways as previously described. Substantially the same procedures described for the identification of the gene for the lymph node homing receptor $gp90^{Mel-14}$ may be employed for the LPAM subunits a_{4m} or b_p.

The DNA may be used to provide conjugates for specific binding to complementary sequences in a host cell. In this way one may identify cells comprising

mRNA for the homing receptor proteins. Furthermore, such sequences may be used as therapeutic agents to destroy expression of homing receptor in cells expressing the homing receptor, by linking such sequences to agents capable of cleaving nucleic acid sequences, such as ribozymes, metal chelates, etc.

The subject proteins may also be used to provide vaccines, by introducing a sequence coding for the subject proteins in place of a gene in a virus encoding the envelope protein. The viruses would then be transported to a site having a large lymphocyte population, where the virus could be endocytosed resulting in a strong immune response.

The subject proteins or fragments thereof may find use as conjugates to various compounds, aggregations, cells or the like, for directing specific compositions to the target site. The epitopic binding site of the homing receptor may be radiolabelled for specifically directing a radioisotope for diagnosis or therapy to high endothelial venules of Peyer's patches or other mucosal sites or lymph nodes. In this way the radiolabel may be concentrated at sites of interest for diagnosis of neoplasia, treatment of aberrant cells, etc. The subject epitopic site may be used for directing cytotoxic compounds to specific sites, such as natural toxins, antibiotics, enzyme inhibitors, or the like. The subject compounds may be bound to liposomes by conventional ways for directing a liposome to a particular site. The lumen of the liposome may serve to carry drugs or other compounds of interest to the site for diagnosis or therapy. Examples of conjugation of proteins to lipids finds extensive exemplification in the literature.

The subject proteins may be used to direct specific subtypes of antibodies or cells producing particular antibodies to target sites, providing protective antibody at the target sites. Thus, IgG,

IgA, IgM, IgE or IgD may find particular use. In addition, the variable regions of antibodies have been cloned and shown to be effective in binding to particular epitopes. By joining a DNA sequence of a peptide capable of homing to a target site with a gene coding the variable region, or if desired, the heavy or light chain of an antibody, fusion protein products may be produced which will provide the desired binding capability at the target site.

The subject nucleic acid gene sequences may also be used to transform cells in order to direct the cells to particular target sites. DNA constructs may be introduced in vitro into a target cell to provide homing capability to the cell. Thus, cells, e.g., lymphocytes, may be transformed with expression casettes comprising a transcriptional and translational initiation region functional in the host cell, a gene encoding one or the other homing receptors or of the subunits of a homing receptor, and a functional translational and transcriptional termination region. It is found that activated lymphocytes lose their homing receptors on the surface. By using an initiation region which is not subject to the natural regulation, the activated lymphocytes would have the homing receptor on the surface and be directed to the target site.

In administering the various therapeutic agents, for the most part, empirical determinations will be involved in the level of the therapeutic agent. The level of proteins which are administered will depend to a substantial degree on the stability of the protein, its size, the manner of administration, the site of administration, the purpose of the therapy, and the like. Therefore, no simple range may be given which would indicate what levels should be applied for any particular therapy. For the most part, the proteins will be administered in an appropriate

physiologically acceptable medium, e.g., water, saline, phosphate buffered saline, or the like. Administration will normally be parenteral, particularly intravascularly. For the reasons given above, the course of treatment will also vary. For therapeutic use of cells, the number of cells will also vary as indicated above.

The subject compositions may be used in diagnostic assays for the proteins or the nucleic acids. Thus the proteins may be used as standards, conjugated to labels as reagents, or the like to determine the presence of the subject protein on a cell. Cells may then be segregated in accordance with their target by using a FACS, the number of cells for a particular target determined as an indication of the health status of an individual, or the like. The nucleic acids may be used as probes to detect transcription of the gene encoding the subject peptides as indicative of the state of the cell, e.g. activated or not activated, the nature of the integrin, or the like. Conventional assay techniques may be used to determine the various events.

The following examples are offered by way of illustration and not by way of limitation.

EXPERIMENTAL

MATERIALS AND METHODS

Cell Lines

The major index cell line utilized for these studies was EL-4/MEL-14^{hi}, a variant of the continuous T cell lymphoma cell line, EL-4, selected by fluorescence flow cytometry for high level expression of the MEL-14 antigen, a property which cosegregated with the capacity to bind peripheral node venules. Additional variants of El-4 with respect to gp^{90}Mel-14 expression, also selected by fluorescence flow

cytometry, were also used in these studies.

Immunoprecipitation and SDS-PAGE analysis of the putative lymphocyte homing receptor, $gp90^{Mel-14}$

Immunoprecipitation of cell surface ^{125}I-iodinated EL4/MEL-14^{hi} by MEL-14 antibody was performed as follows. 2 X 10^7 cells were surface radioiodinated using lactoperoxidase, then solubilized in 2 ml PBS containing 1% Triton X-100, 0.5% deoxycholate, 0.1% SDS, O.1M NaCl, 0.01 M Na phosphate, pH 7.5, and 5mM PMSF according to the method of Witte, et al. Proc. Natl. Acad. Sci. USA 75:2488 (1978) and clarified by ultracentrifugation (30 minutes at 30,000 rpm). The lysate was incubated with a 20X concentrated MEL-14 hybridoma supernatant (Gallatin et al. Nature 304:30-34 (1983)), equivalent to 10-20 µg of monoclonal antibody, for 3-4 hours at 4°C, followed by the addition of a four-fold excess over first stage of affinity purified goat anti-rat IgG, and incubated overnight at 4°C to effect formation of a solid precipitate. The precipitate was centrifuged at 3,000 rpm, and washed three times in 0.01 Tris-HCl, pH 7.4, 0.15 M NaCl, 0.2% Nonidet P40. Remaining complexes were solubilized by heating to 90°C for 3 minutes in Laemmli sample buffer, and analyzed on 10% SDS polyacrylamide tube gels in the Laemmli discontinuous gel system, as modified by Cullen et al. Transplant Rev. 30:236 (1976). The profile was obtained by gel fractionation at 1mm intervals followed by counting.

Immunoprecipitation by MEL-14 antibody of EL4/MEL-14^{hi} cells metabolically labelled with ^3H-Leucine was performed as follows. 2 X 10^8 cells were labeled with 10 mCi of ^3H-Leucine. Briefly, cells were harvested in rapid growth phase and placed in culture at 10^7 cells/ml for 4-6 hours in Spinner balanced saline solution (Gibco), 10% fetal calf serum, supplemented with all amino acids except leucine, which

was added only as isotope at 200 mCi/ml. Cells were washed and solubilized in 0.5% Nonidet P40, 20mM PMSF, for one hour at 4°C. Nuclei and debris were removed by centrifugation, 15 minutes at 3,000 rpm. The lysate was applied to a column of Lens culinaris lectin conjugated to Sepharose 4B equlibrated in 0.01 M Tris-HCl, pH 7.4, 0.15M NaCl, 0.25% Nonidet P40. The glycoprotein enriched pool was eluted with 0.3 M methyl D-mannopyranoside. Precipitation of 5 X 10^6 cell equivalents and SDS-PAGE analysis was as described in the previous paragraph. Gel fractions were incubated in 0.1% SDS overnight at 4°C to elute radioactivity. Radioactivity was counted in a Biofluor scintillation fluid (New England Nuclear) in a Beckman LS counter (Model LS-230). Molecular weight markers: phosphorylase b, 97,400; bovine serum albumin, 68,000; ovalbumin, 43,000.

Two-dimensional polyacrylamide gel analysis of $gp90^{Mel-14}$

3,000 cpm of ^{3}H-phenylalanine labelled $gp90^{MEL-14}$ was dialyzed against distilled water overnight, lyophilized, and solubilized in 20 µl of isoelectric focusing sample buffer. (O'Farrell, Cell 12:1133 (1977)). The first dimensional charge separation was accomplished using non-equilibrium pH gradient electrophoresis (NEPHGE) , which allows assessment of proteins over a broad pH range from 3.5-10. First dimension NEPHGE gels were equilibrated in SDS sample buffer, and for the second dimension were embedded on top of 10% polyacrylamide SDS slab gels and run as described by (O'Farrell, J. Biol Chem. 250: 4007 (1975)). The slab gel was fixed in 40% methanol, 10% acetic acid and stained with Coomassie blue, treated with the fluorographic medium Enhance (New England Nuclear), dried, and exposed to Kodak XAR-5 film for 14 days for autoradiography.

Amino terminal automated Edman degradation of $gp90^{Mel-14}$ intrinsically labelled with ^{3}H-amino acids

Automated sequence analysis was performed on an Applied Biosystems Model 470A gas-liquid phase protein sequenator, modified to bypass the flask for conversion of thiazolinone derivatives. Entire butyl chloride extracts containing the 2-anilinothiazolinone derivatives at each step were transferred into vials directly for scintillation counting in toluene/PPO. Each sample was counted in duplicate for 10 minutes on a Beckman LS counter (Model LS-7500). Positions containing radioactivity above background indicate the presence of a particular ^{3}H-amino acid at that position.

Cyanogen bromide analysis

Automated amino terminal sequencing was performed as described above. CNBr digestion was performed basically as described. Briefly, the glass fiber filter containing the sample is removed from the sequenator, acylated with trifloroacetic anhydride to block remaining free amino groups, and digested by wetting with 25 µl of CNBr solution (100mg/ml in 60% trifloroacetic acid) in a closed container at room temperature for 20 hours. The filter is then returned to the gas phase sequenator for resumption of sequencing analysis.

Endoglycosidase F digestion and SDS-PAGE analysis of cell surface ^{125}I-iodinated $gp90^{Mel-14}$ immunoprecipitates.

2×10^{7} $EL\text{-}4/MEL\text{-}14^{hi}$ cells were labelled with 2 mCi ^{125}I via lactoperoxidase catalyzed surface radioiodination and solubilized and clarified as described. 1 ml aliquots of lysate were incubated with agitation overnight at 4°C with Sepharose 4B conjugated

to MEL-14 antibody (2mg antibody protein/ml gel bed; 50µl conjugated beads per precipitate) or with Sepharose 4B conjugated to R7D4, an isotype matched rat monoclonal antibody negative control which recognizes the immunoglobulin idiotype on 381C13 cells (R. Levy, Stanford University). After four washes with lysis buffer, the samples were eluted with buffer containing 1% SDS, 1% 2-mercaptoethanol, and 1% NP40 by heating to 90°C for 3 minutes. Endoglycosidase F (Endo F) digestions were carried out according to Elder, et al. Proc. Natl. Acad. Sci. USA 79:4s540 (1981). Briefly, eluates were divided into three equal volumes and diluted in reaction buffer to a final concentration of 0.1M Na phosphate, pH 6.1, 0.05 M EDTA, 1% NP40, and 1% 2-mercaptoethanol. The samples were then incubated for 1 to 22 hours at 37°C, either with or without addition of 5µl purified protease free Endoglycosidase F. The reaction was terminated by the addition of SDS to a concentration of 1% and the samples were analysed on a 9% SDS polyacrylamide gel by the method of Laemmli. The gel was dried and fluorographed on Kodak XAR-5 film for 7d.

Immunoprecipitation and SDS-Page analysis of cell surface 125-iodinated E14-MEL-14^{hi} with monoclonal anti-ubiquitin antibodies

1×10^8 EL-4/MEL-14^{hi} cells were labeled with 4 mCi ^{125}I via lactoperoxidase catalyzed radioiodination. Cell viability was assessed at 99%. The cells were lysed at 1×10^7 cells/ml buffer and clarified as described above, except that prior to ultracentrifugation, lysates were precleared by a 4 hour incubation at 4°C with 0.5 ml packed volume of Staphylococcus aureus (IgGsorb, The Enzyme Center, Inc.) and 0.5 ml Sepharose 4B prior to immunoprecipitations. CNBr activated Sepharose 4B (Pharmacia) was conjugated to affinity purified Goat-anti-mouse IgG

(Pelgreeze, 2mg antibody/ml gel bed) and this conjugated material was then incubated for 4 hours at 4°C with ultrafiltration concentrated (5X) preparations of mouse hybridoma supernatant to be used for precipitation (5 ml supernatant equivalent/25µl Sepharose). After thorough washing, antibody-coated Sepharose was incubated overnight in an ice-water bath with labeled cell lysate (25µl Sepharose/ml lysate). Samples were washed four times in lysis buffer, eluted in Laemmli sample buffer, and subjected to SDS-PAGE analysis on a 9% SDS polysacrylamide gel under reducing conditions.

Oligonucleotide synthesis

A 32-fold degenerate fifteen base oligonucleotide corresponding to the amino-terminal five amino acids of the mature protein, determined by single tritiated-amino acid metabolic labeling was synthesized on an Applied Biosystems nucleotide synthesizer as described (Hewick et al. J. Biol. Chem. 256:7990 (1981). Synthesis was performed in four pools of eight-fold degeneracy each, corresponding to the following sequences: 1) 5' TGG AC(T/C) TA(T/C) CA(T/C) TAT 3'; 2) 5' TGG AC(T/C) TA(T/C) CA(T/C) TAC 3'; 3) 5' TGG AC(A/G) TA(T/C) CA(T/C) TAT 3'; 4) 5' TGG AC(A/G) TA(T/C) CA(T/C TAC 3'. These probes were designed to correspond to the RNA coding (sense) strand. In addition, similar probes corresponding to the opposite (anti-sense) strand were also synthesized.

Isolation of RNA

Total RNA was prepared by the guanidine-thiocyanate RNA extraction procedure. Briefly, cells or tissues were homogenized in 8-16 volumes of 5M guanidine-thiocyanate in a Polytron homogenizer. Homogenate was centrifuged 10 min at 10,000 rpm to remove insoluble material. Supernatant was removed to

a new tube and 0.5 volume of 5.7 M CsCl, 100mM EDTA, p
7.5 was added. N-lauryl sarcosine (Sigma) was then
added to 4% w/v. The solution was then carefully
overlayed onto a 5 ml 5.7M CsCl, 0.1M EDTA pH 7.5
cushion in SW polyallomat tubes, 30-40 ml homogenate
per tube. Centriguation was then performed at 20°C,
24,000 rpm for 20 hours in an SW 25 rotor. The
resultant pellet was resuspended in NETS buffer (0.1M
NaCl, 0.001M EDTA, 0.01M Tris-HCl, pH 7.5, 0.2% SDS),
the solution extracted with an equal volume of phenol,
then extracted twice with chloroform.

Alternatively, total cytoplasmic RNA was prepared as previously described, briefly recounted as follows. Cells or tissue homogenates were centrifuged at 1500 rpm 4°C, 10', and resuspended in 20 ml of ice-cold isotonic high pH buffer (IHB) (140mM NaCl, 10mM Tris-HCl, pH 8.4, and 1.5 mM $MgCl_2$). An additional 20 ml IHB, 1.0% NP40 were added and the lysate allowed to sit 5'. Nuclei were centrifuged at 4300 rpm, 10' and supernatant was removed and treated with 1/10 volume Proteinase K, and SDS added to 0.5%. Digestion was allowed to proceed at room temperature for 30'. EDTA was added to a final concentration of 5mM, and the mixture extracted with phenol:chloroform, then again with chloroform, and precipitated with EtOH.

Poly-A containing mRNA was isolated on oligo-dT cellulose as described, from either total or total cytoplasmic RNA as follows. Approximately 0.25g of oligo-dT cellulose were placed in a sterile column and washed with 10 column volume 0.1N NaOH in ETS buffer (1mM EDTA, 10mM Tris-HCl, pH 7.2, 0.25% SDS) and equilibrated in High Salt Buffer (HSB) (0.5M NaCl, 10mM Tris-HCl, pH 7.4, 50mM $MgCl_2$). Total RNA was applied to the column after heating to 65°C for 5', allowed to slowly run through, and the eluate reapplied 3X, with subsequent wash of the column with another 15-20 ml HSB. Bound material was eluted with 4ml ETA, the

solumn was reequilibrated with HSB, and eluted material was reapplied after again heating to 65°C, and adjusting NaCl concentration to 0.5 M, then bound and eluted as described above. Eluted poly-A containing mRNA was ethanol precipitated, resuspended in sterile distilled water, and stored at -70 C.

cDNA synthesis

Poly-dT primed cDNA was synthesized from 4 µg p-A selected mRNA following the basic RNAse H procedure of Gubler and Hoffman. Double-stranded CDNA was modified by placing XhoI adapters on the ends of the cDNA species and the population of cDNA molecules was ligated into the XhoI site of lambdaZAP gt10 vector (Stratagene, Inc.).

Screening of cDNA libraries

Approximately 7.5×10^5 phage plaques in E. coli strain LE 392 were plated onto 150 mm agar plates at about 15,000 plaques per plate, lifted onto nitrocellulose filters, denatured in base, neutralized, and baked for 2 hours at 80°C. Synthetic oligonucleotides were labeled with ^{32}P ATP utilizing polynucleotide kinase. Hybridization was performed in 5X SSPE, 5X Denhardt's, 0.5% SDS, at 25°C. for 18 hours. Filters were subsequently washed in several changes of 5X SSPE, 0.2% SDS. Probing of filters with one pool of oligonucleotides of eight-fold degeneracy, constructed and deduced from the protein sequence obtained (5' TGG AC(A/G) TA(T/C) CA(T/C) TAT 3'), resulted in the identification of 58 independent isolates which reproducibly hybridized with this set of oligonucleotides. These purified clones were excised using helper phage and recircularized to generate sublones in the phagemid vector pBluescript SK(-) (Stratagene, Inc.) for sequence analysis.

DNA sequencing analysis

Fragments of clones or entire clones were sequenced either in the pBluescript SK(-) excised from original lambdaZAP isolates, Bluescript KS(-), or versions of phage M13, mp18 and mp19, modified to include a Not I site for convenient directional cloning. Dideoxy-DNA sequencing was utilized, employing the engineered T7 DNA polymerase Sequenase technique (U.S. Biochemical Corp.). After identification of the clone encoding the protein predicted by radiolabelled amino acid sequencing of $gp90^{Mel-14}$, separate appropriate restriction fragments were subcloned to derive the internal sequence, and thereafter, oligonucleotide sequencing primers were synthesized to obtain the remaining sequence of the full-length clone and to derive second strand sequence where needed.

mRNA blot hybridization analysis

Northern blot analysis was performed on a variety of poly A-selected RNA species isolated from a variety of tissue and cell line sources by the formaldehyde procedure. Approximately 5µg of RNA were applied to each gel lane, and after electrophoresis, RNA was transferred to Genetran nylon filter. Hybridization to isolated insert DNA, labelled with ^{32}P-dCTP hexamer-primed procedure, was performed to 18 hours at 42°C, 50% formamide, 5X Denhardt's, 5X SSPE. Nylon filters were washed at high stringency with rinses of 2X SSPE, 0.2% SDS, room temperature, followed by 0.1X SSPE, 65°C, for 30' 2X. Autoradiographs were developed after exposure to XAR-5 film.

RESULTS

Amino terminal protein sequence analysis of $gp90^{MEL-14}$

The amino terminal protein s quence obtained by automated sequence analysis of material purified

from extracts of MEL-14 positive cells was compared to the protein sequence encoded by the $mLHR_C$ cDNA clone. Purification of $gp90^{MEL-14}$ from EL-4/MEL-14hi cells metabolically labelled with radiolabelled amino acids was performed as described ([24]) using the monoclonal MEL-14 antibody. 2 X 10^8 cells were labelled with 10 mCi of a single ^3H- or ^{35}S-amino acid ([84]) for 4-6 hours in Spinner balanced salt solution (Gibco), 10% fetal calf serum was supplemented with all amino acids except the radiolabelled one which was added at 200 Ci/ml. Cells were solubilized in 0.5% Nonidet P40, 20mM PMSF, for one hour at 4°C. Nuclei and debris were removed by centrifugation for 15 minutes at 3,000 rpm. The lysate was applied to a column of Lens culinaris lectin conjugated to Sepharose 4B equilibrated in 0.01 M Tris-HCl, pH 7.4, 0.15M NaCl, 0.25% Nonidet P40. The column was washed and bound material eluted with 0.3 M methyl-D-mannopyranoside. Lysates were incubated with a 20X concentrated MEL-14 hybridoma supernatant, equivalent to 10-20 micrograms of monoclonal antibody, for 3-4 hours at 4°C, followed by the addition of an excess of affinity purified goat anti-rat IgG and incubation overnight at 4°C. The precipitate was centrifuged at 3,000 rpm, and washed three times in 0.01M Tris-HCl, pH 7.4, 0.15M NaCl, 0.25% Nonidet P40. Complexes were solubilized by heating to 90°C for three minutes in Laemmli sample buffer, and $gp90^{MEL-14}$ was purified on 10% SDS polyacrylamide tube gels using the Laemmli discontinuous gel system ([85]), as modified by Cullen ([86]). Gel fractions were incubated in 0.1% SDS overnight at 4°C to elute the protein. Radiolabelled fractions were monitored in Biofluor scintillation fluid (New England Nuclear) in a Beckman LS counter (Model LS-230).

Automated sequence analysis was performed on an Applied Biosystems Model 470A gas-liquid phase protein sequenator, modified to bypass the flask for

conversion of thiazolinone derivatives, as described ([24]). Entire butyl chloride extracts containing the 2-anilino-thiazolinone derivatives at each step were transferred into vials directly for scintillation counting in toluene/PPO. Each sample was counted in duplicate for 10 minutes on a Beckman LS counter (Mod LS-7500). Positions containing radioactivity above background indicated the presence of a particular ^{3}H- or ^{35}S-amino acid.

Unambiguous amino acid assignments could be made at 15 of the amino terminal 32 residues, and an additional tentative tyrosine position was made at residue 37.

mRNA blot hybridization analysis of MEL-14 positive and negative tissues and cell lines

The index cell line utilized for these studies was EL-4/MEL-14hi, a variant of the continuous T-c ll lymphoma cell line, EL-4, selected by fluorescence activated flow cytometry for high level expression of the MEL-14 antigen, a property which cosegregated with the capacity to bind peripheral node venules. Additional variants of EL-4, differing with respect to $gp90^{MEL-14}$ expression were obtained from various sources, following fluorescence activated cell sorter (FACS) analysis, and used in these studies as well. C6V1 and VL3 are both radiation-induced leukemia virus thymoma clonal cell lines. Northern blot analysis was performed by the formaldehyde procedure as described ([88]), on a variety of poly A-selected RNA species isolated from a variety of tissues and cell lines. Approximately 5 µg of RNA were applied to each gel lane, and after electrophoresis RNA was transferred to Genetran nylon filter. Hybridization to probe labelled with ^{32}P-dCTP using the random primer procedure ([89]) was performed for 18 hrs at 42°C, 50% formamide, 5X Denhardt's, 5X SSPE. Nylon filters were washed with 2X

SSPE, 0.2% SDS, at room temperature, followed by 0.1X SSPE, 65°C, for 30', twice. Autoradiographs were developed after X hour exposure to XAR-5 film. a). Hybridization using ^{32}P-labelled murine lymph node homing receptor core peptide (mLHRc) DNA. Lane A, EL-4/MEL-14xhi; lane B, el-4/MEL-14hi; lane C, BD EL-4/MEL-14hi; lane D, EL-4/MEL-14lo; lane E, BD EL-4/MEL-14lo; lane F, VL3; lane G, C6V1; lane H, thymus; lane I, spleen; lane J, mesenteric lymph node; lane K, liver; lane L, kidney; lane M, testes; lane N, brain; b) Hybridization using ^{32}P-labelled actin. Lanes A-N as in a).

All cell lines and tissues expressing detectable nRNA exhibited identical patterns with bands at 1.5, 2.5 and 5.2 kb. Intensity of hybridization correlated with the cell surface expression of gp90^{MEL-14}.

A number of variants of EL-4 which differ with respect to level of expression of gp90^{MEL-14} were selected and sorted by fluorescence flow cytometry, then quickly grown to process mRNA. The patterns on Northern blot analysis corresponding to mRNA from EL-4/MEL-14Xhi, EL-4/MEL-14hi and EL-4/MEL-14lo show transcript abundance paralleling cell surface expression. A particular decrement in the amount of 1.5 kb transcript species in the EL-4/MEL-14lo cells relative to the other positive cell lines suggested a prominant role for this species in determining cell surface expression.

Additional cell lines unrelated to EL-4 were also included in the analysis, VL-3 and C6VL, <u>in vitro</u> T-cell lymphoma lines, the former expressing relatively low levels of surface antigen and the latter showing no cell surface staining. The transcript pattern paralleled surface expression, to roughly control for relative loading of mRNA on the gel, the filter was stripped and rehybridized with sequences of a

relatively ubiquitous transcript, the beta-actin gene. Hybridization was reasonably homogenous between lanes, indicating that the differences observed for the transcript were related to abundance.

Distribution in normal tissues shows a predominant lymphoid distribution, paralleling tissue staining patterns for MEL-14. Thymus, spleen and masenteric lymph nodes are positive for the same size transcript found in cell lines, while liver, kidney and brain show no detectable transcripts.

Fluorescence activated cell sorter (FACS) analysis of cell lines varying with respect to expression of $gp90^{MEL-14}$

Cells were stained with an isotype matched control (A1, B1, C1, D1) or MEL-14 hybridoma supernatant (all others), followed by FITC-conjugated goat anti-rat immunoglobulin absorbed for cross reactivity with mouse serum components. A). 1: EL-4hi; 2: EL-4lo; 3: EL-4hi; 4: EL-4Xhi. B). 1: BD EL-4lo; 2: BD EL-4lo. C). 1: VL3; 2: C5V1; 3: VL3. D). 1: BD EL-4lo, positive sort; 2: BD EL-4llo, negative sort; 3: BD EL-4lo, positive sort.

An independent EL-4 clonal cell line was identified which demonstrated a distinct MEL-14 staining pattern containing two discrete populations of cell expression - a predominant negative population and a relatively small population, about 5% of cells expressing $gp90^{MEL-14}$. The 3% highest and lowest intensity staining cells were sorted, immediately grown, and mRNA extracted. Expression of the transcript in Northern blot is present in the high population and absent in the negative population, thereby showing, in combination with the variants described above, cosegregation of transcript and cell surface antigen expression in variants derived from the same clonal cell line.

FACS analysis of Cos-7 cells transfected with $mLHR_C$ DNA

The full length cDNA clone was transferred to the expression vector CDM8, a plasma with tetracycline/ampicillin resistance containing CMV (cytomegalo virus)/HIV (human immunodeficiency virus) promoters, SV40, and M13 origins of replication, splice and polyadenylation sites, and a polylinker region for insertion of cDNA species (Brian Seed). Plasmid DNA was transfected into confluent Cos-7 cells using the DEAE-dextran transfection procedure as described (^{35}). Enrichment of MEL-14 positive transfectants was achieved by planting transfected cells stained with MEL-14 onto goat anti-rat Ig coated petri dishes. Non-adherent cells were removed, and after 0.5-1 hr, adherent cells were reanalysed by fluorescence staining.

The results of analysis of the transfected cells show a population of positive cells when stained with MEL-14 compared to staining with an isotype matched control antibody. Identical backgrounds were obtained staining mock transfected or Thy-1 transfected Cos-7 cells with MEL-14.

Immunoprecipitation of MEL-14 reactive cell surface determinant(s) from enriched $mLHR_C$ transfected Cos-7 cells

$mLHR_C$ transfected Cos-7 cells enriched as described above were surface labelled with ^{125}I using lactoperoxidase (^{90}). Immunoprecipitations and electrophoreses were performed as described above using slab polyacrylamide gels under non-reducing and reducing conditions. Non reducing gel: A: transfectants, isotype control; B: transfectants, MEL-14 antibody; C: EL-4/MEL-14hi, MEL-14 antibody. Reducing gel: A: transfectants, MEL-14 antibody; B: transfectants, isotype control; C: EL-4/MEL-14hi,

MEL-14 antibody.

The results demonstrated the presence of 2 MEL-14 specific species under non-reducing conditions, one slightly smaller than the mature GP-90^{MEL-14} from EL-14/MEL-14hi (lane C) with an apparent molecular weight of slightly less than 70 kD and an even smaller band of about 60 kD. This indicates that there may be processing of the transcript into discrete forms and perhaps reflects alternative pathways of post-translational modification, including glycosylation and/or ubiquitination. It should be noted that, while the molecular sizes are slightly altered from the EL-4/MEL-14hi form, the typical reducing/non-reducing behavior of GP90^{MEL-14}, with non-reduced form migrating faster than reduced form, is retained.

Complete nucleotide and predicted protein sequence of mLHR$_C$ cDNA

The nucleotide sequence of the cDNA was determined by the dideoxy chain termination method of Sanger and Coulsen, employing the engineered T7 DNA polymerase Sequenase system (U.S. Biochemical Corp.). Single stranded template DNA's were derived from either pBluescript SK(-) (excised from original lambdaZAP isolates), Bluescript KS(-), or versions of bacteriophage M13mp18 and mp19 modified to include a Not I site for convenient directional cloning. Once sequences encoding the amino terminus predicted by amino acid sequencing gp90^{MEL-14} were identified, appropriate restriction fragments were subcloned to derive the internal sequence. Subsequently, oligonucleotide primers were synthesized to obtain the remaining sequence of the full-length clone and to obtain second strand sequence where needed. The predicted protein sequence is indicated below beginning with the initiator methionine at nucleotide position 54; numbering to the right indicates the nucleotide and

protein positions. Cysteine residues in the mature protein are marked with an asterisk (*) above, and canonical N-linked carbohydrate recognition sites (Asn-X-Ser/Thr) are overlined with arrow bars. The 15 nucleotides encoding the amino terminal five amino acids and hybridizing to the oligonucleotide probe used for screening are underlined in bold. Poly-A splice and common polyadenylation recognition sequences are double underscored.

```
CA GGT GGA GGA GGC TGA GGC TGC AGA GAG ACT TGC AGA GAG ACC CAG CAA GCC Met Val Phe Pro Trp Arg Cys Glu Gly Thr Tyr Trp  -27
                                                                     ATG GTG TTT CCA TGG AGA TGT GAG GGT ACT TAC TGG   89
Gly Ser Arg Asn Ile Leu Lys Leu Trp Val Trp Thr Leu Leu Cys Cys Asp Phe Leu Ile His His Gly Thr His Cys Trp Thr Tyr His
                                                                                                   -1  +1
GGC TCG AGG AAC ATC CTG AAG CTG TGG GTC TGG ACA CTG CTC TGT TGT GAC TTC CTG ATA CAC CAT GGA ACT CAC TGT TGG ACT TAC CAT    4
Tyr Ser Glu Lys Pro Met Asn Trp Glu Asn Ala Arg Lys Phe Cys Lys Gln Asn Tyr Thr Asp Leu Val Ala Ile Gln Asn Lys Arg Glu  179
TAT TCT GAA AAG CCC ATG AAC TGG GAA AAT GCT AGA AAG TTC TGC AAG CAA AAT TAC ACA GAT TTA GTC GCC ATA CAA AAC AAG AGA GAA   34
Ile Glu Tyr Leu Glu Asn Thr Leu Pro Lys Ser Pro Tyr Tyr Tyr Trp Ile Gly Ile Arg Lys Ile Gly Lys Met Trp Thr Trp Val Gly  269
ATT GAG TAT TTA GAG AAT ACA TTG CCC AAA AGC CCT TAT TAC TAC TGG ATA GGA ATC AGG AAA ATT GGG AAA ATG TGG ACA TGG GTG GGA   64
Thr Asn Lys Thr Leu Thr Lys Glu Ala Glu Asn Trp Gly Ala Gly Glu Pro Asn Asn Lys Lys Ser Lys Glu Asp Cys Val Glu Ile Tyr  359
ACC AAC AAA ACT CTC ACT AAA GAA GCA GAG AAC TGG GGT GCT GGG GAG CCC AAC AAC AAG AAG TCC AAG GAG GAC TGT GTG GAG ATC TAT   94
Ile Lys Arg Glu Arg Asp Ser Gly Lys Trp Asn Asp Asp Ala Cys His Lys Arg Lys Ala Ala Leu Cys Tyr Thr Ala Ser Cys Gln Pro  449
ATC AAG AGG GAA CGA GAC TCT GGG AAA TGG AAC GAT GAC GCC TGT CAC AAA CGA AAG GCA GCT CTC TGC TAC ACA GCC TCT TGC CAG CCA  124
Gly Ser Cys Asn Gly Arg Gly Glu Cys Val Glu Thr Ile Asn Asn His Thr Cys Ile Cys Asp Ala Gly Tyr Tyr Gly Pro Gln Cys Gln  539
GGG TCT TGC AAT GGC CGT GGA GAA TGT GTG GAA ACT ATC AAC AAT CAC ACG TGC ATC TGT GAT GCA GGG TAT TAC GGG CCC CAG TGT CAG  154
Tyr Val Val Gln Cys Glu Pro Leu Glu Ala Pro Glu Leu Gly Thr Met Asp Cys Ile His Pro Leu Gly Asn Phe Ser Phe Gln Ser Lys  629
TAT GTG GTC CAG TGT GAG CCT TTG GAG GCC CCT GAG TTG GGT ACC ATG GAC TGC ATC CAC CCC TTG GGA AAC TTC AGC TTC CAG TCC AAG  184
Cys Ala Phe Asn Cys Ser Glu Gly Arg Glu Leu Leu Gly Thr Ala Glu Thr Gln Cys Gly Ala Ser Gly Asn Trp Ser Ser Pro Glu Pro  719
TGT GCT TTC AAC TGT TCT GAG GGA AGA GAG CTA CTT GGG ACT GCA GAA ACA CAG TGT GGA GCA TCT GGA AAC TGG TCA TCT CCA GAG CCA  214
Ile Cys Gln Val Val Gln Cys Glu Pro Leu Glu Ala Pro Glu Leu Gly Thr Met Asp Cys Ile His Pro Leu Gly Asn Phe Ser Phe Gln  809
ATC TGC CAA GTG GTC CAG TGT GAG CCT TTG GAG GCC CCT GAG TTG GGT ACC ATG GAC TGC ATC CAC CCC TTG GGA AAC TTC AGC TTC CAG  244
Ser Lys Cys Ala Phe Asn Cys Ser Glu Gly Arg Glu Leu Leu Gly Thr Ala Glu Thr Gln Cys Gly Ala Ser Gly Asn Trp Ser Ser Pro  899
TCC AAG TGT GCT TTC AAC TGT TCT GAG GGA AGA GAG CTA CTT GGG ACT GCA GAA ACA CAG TGT GGA GCA TCT GGA AAC TGG TCA TCT CCA  274
Glu Pro Ile Cys Gln Glu Thr Asn Arg Ser Phe Ser Lys Ile Lys Glu Gly Asp Tyr Asn Pro Leu Phe Ile Pro Val Ala Val Met Val  989
GAG CCA ATC TGC CAA GAG ACA AAC AGA AGT TTC TCA AAG ATC AAA GAA GGT GAC TAC AAC CCC CTC TTC ATT CCT GTA GCC GTC ATG GTC  304
Thr Ala Phe Ser Gly Leu Ala Phe Leu Ile Trp Leu Ala Arg Arg Leu Lys Lys Gly Lys Lys Ser Gln Glu Arg Met Asp Asp Pro Tyr 1079
ACC GCA TTC TCG GGG CTG GCA TTT CTC ATT TGG CTG GCA AGG CGG TTA AAA AAA GGC AAG AAA TCT CAA GAA AGG ATG GAT GAT CCA TAC  334
TGA TTC ATC CTT TGT GAA AGG AAA GCC ATG AAG TGC TAA AGA CAA AAC ATT GGA AAA TAA CGT CAA GTC CTC CCG TGA AGA TTT TAC ACG 1169
CAG GCA TCT CCC ACA TTA GAG ATG CAG TGT TTG CTC AAC GAA TCT GGA AGG ATT TCT TCA TGA CCA ACA GCT CCT CCT AAT TTC CCC TCG 1259
CTC ATT CAT CCC ATT AAC CCT ATC CCA TAA TGT GTG TCT ATA CAG AGT AGT ATT TTA TCA TCT TTT CTG TGG AGG AAC AAG CAA AAG TGT 1349
TAC TGT AGA ATA TAA AGA CAG CTG CTT TTA CTC TTT CCT AAA AAA AAA AAA AAA AAA 1439
```

The cDNA clone has a 54 bp 5'untranslated region followed by an initiator ATG codon, which begins an uninterrupted open reading frame of 1,116 bp. The TGA stop codon at position 1169 is followed by 327 bp of 3' untranslated region.

The reading frame encodes a protein with a hydrophobic leader sequence 38 amino acids in length before reaching the initial tryptophan residue of the mature protein. Hydropathy analysis confirms a generally hydrophobic leader sequence, where the initial 15 residues are neutral to slightly hydrophilic. The signal sequence includes 3 positively charged residues, 4 cysteine residues, and 3 histidine residues, clustered in the 12 residues preceding the mature protein.

The mature protein possesses 10 potential asparagine-linked glycosylation sites, with 6 of these contained within an identical repeat unit structure. The mature protein contains 22 cysteine residues, where 12 of the cysteines are present in the complement regulatory protein repeat structures, and an additional 9 cysteines are concentrated in the 60 amino acids preceding the repeat units involving the EGF-like domain, resulting in a highly cysteine-rich pre-transmembrane region of 180-190 amino acids.

The deduced mature protein is 334 amino acids in length with a calculated molecular weight of 37,600. A hydrophobic transmembrane region encompassing amino acids from about 295-317 is followed by a cluster of positively charged residues and a hydrophilic cytoplasmic tail of 18 amino acids. A hydropathy plot shows distinct regions of relative hydrophilicity, concentrated in the amino terminal 150 amino acids and in the membrane proximal approximate 20 amino acids. The intervening extracytoplasmic portion is comprised of a relatively electrically neutral stretch which includes the presence of the

aforementioned repeat units, identical at both the nucleotide and protein level.

Protein comparisons reveal the extracytoplasmic portion of the receptor to be made of 3 separate extracytoplasmic domains, defined by their homology to 3 disparate protein motifs.

The amino-terminal domain shows homology to the carbohydrate binding domains of animal lectins (position 74-118); the succeeding 37 amino acids (positions 119-155) occupy the region between the lectin domain and the complement regulatory repeat units, exhibit similarity to the epidermal growth factor (EGF) cysteine-rich repeat unit; and the third region is comprised of 2 identical repeat units comforming to the consensus sequence of homologous repeats found in complement regulatory and other proteins (positions 156-217).

The $mLHR_C$ is homologous over a stretch of 45 amino acids equivalent to the 50 carboxy-terminal residues of the binding domain in animal lectins. The region includes three invariant cysteines at 90, 109, and 116 in $mLHR_C$ and -W at 75-76, a characteristic E-T-N (80-82), an E at 88, C-V at 90-91, and the conserved G-WND at 102-106. Only a highly conserved G, position 12 in the consensus sequence and present in other mammalian lectins, is absent from the carbohydrate-binding domain in $mLHR_C$. Between conserved residues N-82 and E-88 there is a cluster of 3 lysine residues and an insertion relative to the consensus sequence of 5 charged amino acids between C-17 and G-24 of the consensus sequence. The entire domain contains 10 positively charged residues, 3 R and 7 K. The presence of a lectin domain in $mLHR_C$ is consistent with studies which have demonstrated that mannose-6-phosphate and some analogs, but not other carbohydrates, inhibit binding to peripheral lymph node HEV, but not Peyer's patch. The unusual Lysine enrichment in the lectin

domain, combined with the known role of this domain in binding and our understanding of the known role of this MEL-14 epitope, suggests this r gion may contain the site of ubiquitination. HEV addressin is inactivated by treatment with neuraminidase, but not alkaline phosphatase, and an as yet unidentified, non-phosphorylated sialic-acid dependent molecule is indicated as the ligand for $mLHR_c$.

The EGF-like domain in $mLHR_c$ consists of a single copy homolog of the EGF repeat unit, which preserves many of the C/G residues characteristic of the structure. All 6 consensus C's are present as well as G's at 147 and 150, and tyrosine at 148 of $mLHR_c$. The relationship of these conserved residues is identical to that of human and bovine blood clotting factors IX and X and the Drosophila Notch gene product (and in all but 4 of 36 repeats in this gene), but not to the other molecules containing EGF-like domains, with no insertions or deletions required to align the sequences. The EGF-like domain shares homology with a portion of one of the cysteine-rich repeat units of the beta chain of the integrin LFA-1β_2 chain in the human (positions 449-483). A 12 amino acid region comprising $mLHR_c$ 142-154 aligns directly with 480-492 of the LFA-1β_2 subunit, retaining the conserved spacing of 3 cysteines, with identity of 7 residues.

The next domain is a precisely duplicated repeat unit, with each unit of 62 amino acids in length, spanning positions 156-217 and 218-279. Murine complement factor H, a serum protein with complement regulatory activity, exhibits significant homology. In factor H, there are 20 contiguous, homologous, though not identical, repeat units having approximately 10-31% homology with the $mLHR_c$ receptor. The same homologous repeat motif exists in a number of complement regulatory proteins which bind C3/C4, and in other proteins such as the Il-2 receptor, the β_2-glycoprotein

serum protein, and factor XIII. The consensus sequence position is represented in the homing receptor repeat unit sequence T-4, P-7, F-30, C-32, G-35, C-46, G-50, W-52, P-57, and C-59. In addition, except for a relative deletion of 1 residue between C-4 and P-7 the consensus sequence and an insertion of 3 residues between P-7 and F-30, relative spacing of the remaining residues of the consensus sequence is completely preserved in the homing receptor sequence.

Homology motifs found in the $mLHR_C$ protein sequence

Proteins having homology were aligned as shown below. The top line in each panel depicts the amino acid residues whose positioning defines the consensus sequence for the particular motif. Residues in parentheses may or may not be present in a sequence conforming to the motif. Dashes indicate positions that must be occupied by an amino acid, while spaces demarcate regions of variable length. A. $mLHR_C$ residues 74-118 compared to the consensus motif for carbohydrate binding domains of animal lectins and representative proteins exhibiting that motif. B. $mLHR_C$ residues 122-155 compared to the consensus motif for cysteine-rich EGF-like repeat units and representative proteins exhibiting that motif. C. $mLHR_C$ residues 156-217 compared to the consensus motif for the complement regulatory repeat units and representative proteins exhibiting that motif. R-MBP-C, rat mannose binding protein C; R-MBP-A, rat mannose binding protein A; H-MBP-H, human mannose binding protein H; CPSa, canine pulmonary surfactant a; RASGPR, rat asialoglycoprotein receptor; HASGPR, human asialoglycoprotein receptor; HFc_eR, human Fc epsilon receptor; CHL, chicken hepatic lectin; ISL, sarcophaga peregrina hemolymph lectin; Ech, echinoidin, lectin from sea urchin coelemic fluid; EGF, epidermal growth factor; TGF, transforming growth factor; tPAhu, human

tissue plasminogen activator; LDL, low density lipoprotein; CR1, complement receptor 1; H, factor H; C_4bp, C_4 binding protein; Ba, factor Ba; βGPI, β-glycoprotein I; Il-2R, interleukin-2 receptor.

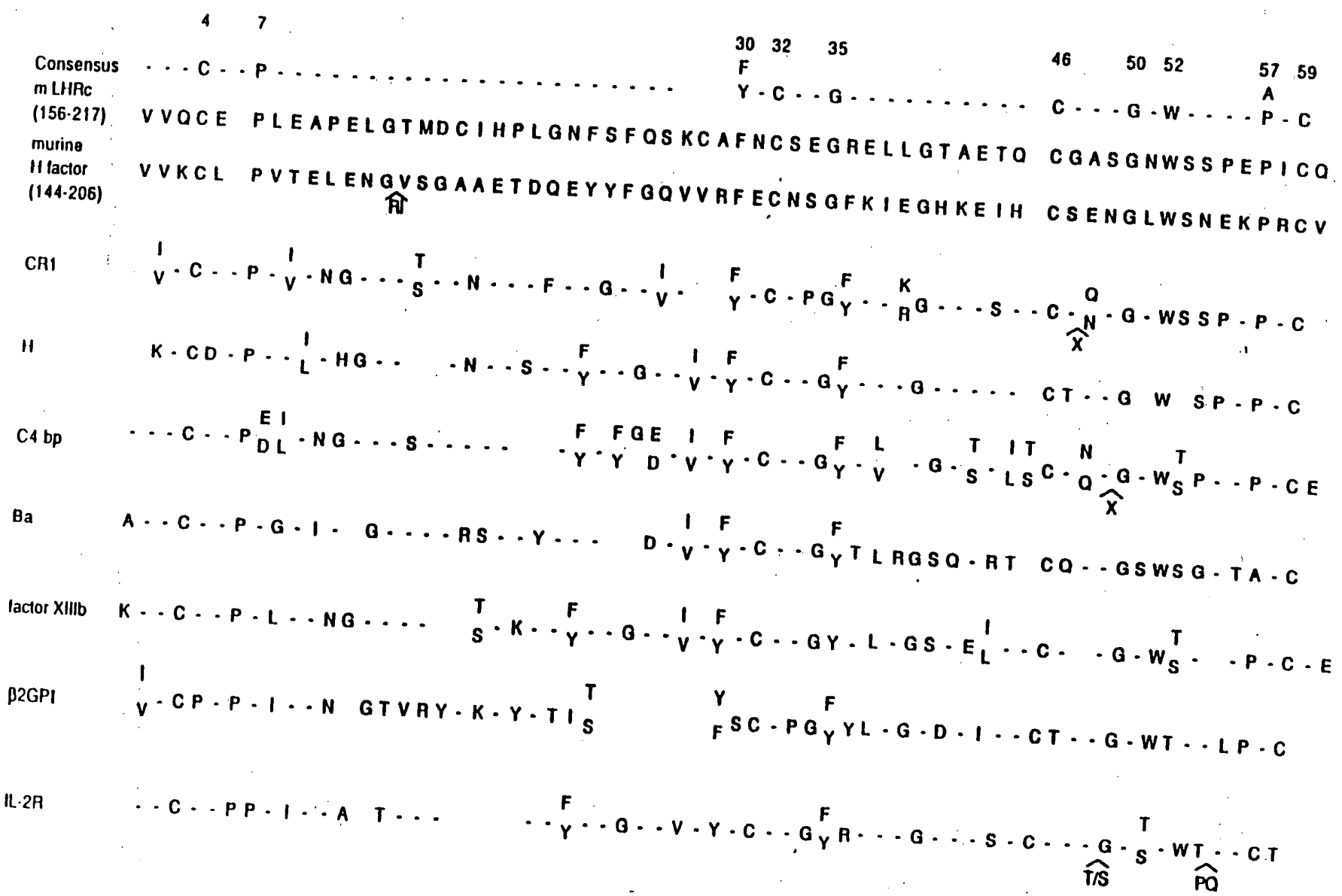
A.

Consensus	-(N)W- - -(E)P(N) - -G(S)- E - C V - - - (N)G - W N D - - C - - - - - - - - C E
mLIIRc (74-118)	E N W G A G E P N N K K S K E D C V E I Y I K R E R D S G K W N D D A C H K R K A A L C Y
RMBPC	T N W N E G E P N N V G S G E N C V V L L T N G K W N D V P C S D S F L V V C E
R-MBP-A	S N W K K D E P N D H G S G E D C V T I V D N G L W N D I S C Q A S H T A V C E
H-MBP-H	Y N W N E G E P N N A G S D E H C V L L L K N G Q W N D S P C I H L P S A V C E
CPSa	T N W Y P G E P R G R G K E Q C V E M Y T D G Q W N N K N C L Q Y R L A I C E
RASGPR	K N W R P G Q P D D W Y G HGLGGG EDC A H F T T D G H W N D D V C R R P W R W V C E
HASGPR	K N W R P E Q P D D W Y G HGLGGG EDC A H F T T D G H W N D D V C Q R P Y R W V C E
$HFc_\beta R$	S N W A P G E P T S R S Q G E D C V M M P Q S G R W N D A F C D R K L G A W V C D
CHL	T F W K E G E P N N R G F N E D C A H V W T S Q Q W N D V Y C T Y E C Y Y V C E
ISL	A Y W S E N N P D N Y K H Q E H C V H I W D T K P L Y E W N D N D C N V K M G Y I C E
Ech	T A W V G S N P D N Y G S G E D C T Q M V M G A G L N W I D L P C S S T R H Y I C K L

B.

Consensus	C · · · · · · · C · · · · · C · · · · · · · · · C · C · GY · G · · C
mLHRc (122-155)	CQPGS CNGRGECVETINNHT CICDAGYYGPQCQY
D. Notch r26	CTESS CLNGGSCIDGINGYN CSCLAGYSGANCQY
D. Notch r14	CQSQP CRNRGICHDSIAGYS CECPPGYTGTSCE
C. eleg. lin-12	CLENP CSNGGVCHQHRESFS CDCPPGFYGNGCEQ
Factor IX prec hu	CESNP CLNGGSCKDDINSYE CWCPFGFEGKKCEL
Factor IX bov	CESNP CLNGGMCKDDINSYE CWCQAGFEGTNCEL
Factor X hu	CETSP CQNQGKCKDGLGEYT CTCLLGFEGKNCEL
Factor X bov	CEGHP CLWQGHCKDGIGDYT CTCAEGFEGKNCEF
EGF-prec mouse	CGPGG CGSHARCVSDGETAE CQCLKGFA^RD GNLCSD
EGF-mouse	CPSSYDGYCLNGGVCMHIESLDSYTCNCVIGYSGDRCQT
EGF-hu	CPLSHDGYCLHDGVCMYIEALDKYACNCVVGYIGERCQY
TGF-rat	CPDSHTQYCFHGT CRFLVQEEKPACVCHSGYVGVRCEH
Protein C bov	CDLP CCGRGKCIDGLGGFR CPCAEGWEGRFCLH
tPA hu	CSEPR CFNGGTCQQALYFSDFVCQCPEGFAGKCCEI
LDL receptor hu	CLDNNGG CSHV ICNDLKIGYE CLCPDGQL^VA QRRCED
compl C9 hu	CHT CQNGGTVILMDGKCL CACPFKFEGIACEI
vaccinia 19 K	CGPEGDGYCLHGD CIHARDIDGMYCRCSHGYTGIRCQH
D. Notch consensus	CXSXP CXNGGTCXDXXXXFX^Y CXCXXGYXGXXCEX

C.


4 7
30 32 35
46 50 52 57 59
Consensus
m LHRc (156-217)
murine H factor (144-206)
CR1
H
C4 bp
Ba
factor XIIIb
β2GPI
IL-2R

cDNA synthesis and screening of human cDNA libraries

Poly-dT primed cDNA was synthesized from 4 µg p-A selected mRNA following the basic RNase H procedure of Gubler and Hoffman. Double-stranded cDNA was synthesized and ligated into the EcoRI site of lambda gt11. Approximately 1.0 x 10^6 phage plaques in E. coli strain LE 392 were plated onto 150 mm agar plates at about 20,000 plaques per plate, lifted onto nitrocellulose filters in duplicate, denatured in base, neutralized, and baked for 2 hours at 80°C. The full-length mouse lymph node homing receptor cDNA clone ($mLHR_C$) was excised as a NotI/NotI restriction fragment of about 1500 bp or as an approximate 1200 bp fragment excised with XhoI emcompassing all but the 5' 300 bp of the full-length clone. These inserts were purified and labeled with 32P alpha-dCTP by the standard hexamer priming method. Hybridization was performed in 5X SSPE, 5X Denhardt's, 0.5% SDS, at 25°C for 18 hours, with duplicate filters, with 45% formamide in one probe mixture and 35% formamide in the other. The NotI full-length probe was placed in the 45% formamide set and the XhoI excised probe in the 35% set. Filters were hybridized for 18 hours and subsequently washed in several changes of 5X SSPE, 0.2% SDS at room temperature, and then at 55°C. Probing of filters resulted in the identification of 8 independent isolates which hybridized in both sets of filters and reproducibly hybridized on rescreening to plaque purification.

Lambda gt11 inserts were isolated and subcloned into the EcoRI site of M13mp19 for sequence analysis by the dideoxy-sequencing method described above.

Human Lymph Node Homing Receptor Sequence

```
GAGTGCAGTCTAGGTGCAGCACAGCACACTCCCTTTGGCAAGGACCTGAGACCCT
TGTGCTAAGTAAGAGGCTCAATGGGCTGCAGAAGAACTAGAGAAGGACCAAGCAA
AGCC ATG ATA TTT CCA TGG AAA TGT CAG AGC ACC CAG AGG
     M   I   F   P   W   K   C   Q   S   T   Q   R

ACT TAT GGA ACA TCT TTC AAG TTG TGG GGC TGG ACA ATG CTC
 T   L   G   T   S   F   K   L   W   G   W   T   M   L

TGT TGG GAT TTC CTG GCA CAT CAT GGA ACC GAC TGC TGG ACT
 C   C   D   F   L   A   H   H   G   T   D   C   W   T

TAC CAT TAT TCT GAA AAA CCC ATG AAC TGG
 Y   H   Y   S   E   K   P   M   N   W
```

The following describes the experimental procedures for identifying LPAM-1 and -2.

Antibodies and cell lines

The production of rat monoclonal antibody R1-2 (IgG2b) recognizing the α chain of the LPAM-1 molecule was prepared as follows. Spleen cells from Fisher rats immunized 3X i.p. with the Peyer's patch HEV binding lymphoma line TK1 were fused with a non-secreting mouse myeloma P3x63AG8.653 using standard procedures (Galfre et al (1977) Nature 266:550-552). Hybridomas producing antibodies reactive in immunofluorescence assays with TK1 cells but not HEV binding lymphoma TK5 were cloned by limiting dilution. Cultured supernatants of sub-clones were screened for inhibition of lymphocyte binding to HEV of either peripheral nodes or Peyer's patches. Monoclonal antibody R1-2 (IgG2b, κ) which recognizes the LPAM-1 molecule was chosen for further analysis.

The rat monoclonal antibodies M1/70 (IgG2b)

reacting with the α submit of the murine Mac-1 antigen (Springer et al, (1978) Immunol. 8:539551) and 30G12 (IgG2a) specific for mouse leucocyte common antigen T200 (Ledbetter and Herzenberg, (1979) Immunol. Rev. 47:63-90) were used as controls. Hybridomas M17/4.3 and M18/20 secreting rat monoclonal antibodies specific for the α and β chain of the murine LFA-1 antigen were obtained from Dr. T.A. Springer, Dana-Farber Cancer Institute, Boston. A polyvalent rabbit antiserum raised against a synthetic peptide corresponding to the COOH-terminal domain of the chicken integrin β_1 subunit was obtained from Drs. E.E. Marcantonio and R.O. Hynes, Massachusetts Institute of Technology, Cambridge. This anti-β_1-peptide antiserum was shown to be monospecific for integrin β_1 and reacts with β chains from a variety of vertebrates (Marcantonio and Hynes (1988) J. Cell Biol. 106:1765-1772). The rabbit anti-VLA-β antiserum was obtained from Dr. M.E. Hemler, Dana-Farber Cancer Institute, Boston. The polyvalent rabbit antiserum specific for platelet glycoprotein IIIa (Leung et al., (1981) J. Biol. Chem. 256:1994-1997 was obtained from Dr. L.L.K. Leung, Stanford University, Medical School.

T cell lymphomas TK23, TK40, and TK50 were passaged by subcutaneous injections of 10^4 - 10^7 cells into syngeneic AKR/cum recipients. All other cell lines were maintained in tissue culture using RPMI 1640 with 7% fetal calf serum.

In vitro HEV binding assay

This technique has been described previously in detail (Stamper and Woodruff (1976) J. Exp. Med. 144:828-833; Butcher et al, (1979) J. Immunol. 123:1996-2003). Briefly, lymphocytes in Hank's balanced salt solution (HBSS) containing 5% calf serum and 20mM HEPES pH7.4 were incubated with mild rotation for 30 minutes at 7°C on freshly cut frozen sections of murine peripheral (axillary, brachial, inguinal and

cervical) nodes or Peyer's patches. After incubation, adherent cells were fixed to the tissue section in cold 1.25x PBS containing 2% formaldehyde (J.T. Baker Chemical Co., Phillipsburg, N.J.). After fixation, nonadherent cells were rinsed off with a gentle stream of PBS and the dried sections were examined microscopically. To facilitate quantitative comparisons, an internal standard population of mouse mesenteric node lymphocytes labeled by a 15 minute incubation at 37°C with 40µg/ml fluoresceinisothiocyanate (FITC, Sigma Chemical Co., St. Louis, MO) in serum-free HBSS containing 20mM HEPES pH7.4 was mixed with each sample population before incubation. Lymphocytes adherent to HEV were first selected under darkfield illumination and then scored as sample (unlabeled) or standard (fluorescent) cells with UV epi-illumination. At least twelve sections per experiment were analyzed. The ratio of sample to standard cells on HEV (R_{HEV}) and in the incubation mixture (R_I) was determined and the specific adherence ratio, R_{HEV}/R_I, was calculated for each sample (SAR_s) and for mesenteric node lymphocytes (SAR_m). Direct comparison of the adhesive capacity of sample cells to that of unlabeled mesenteric node lymphocytes is given as a relative adherence ratio ($RAR_s=SAR_s/SAR_m$). The RAR represents therefore the calculated number of sample cells bound to HEV per reference mesenteric node lymphocyte bound under the same conditions.

Cell labeling and immunoprecipitation

Cells were surface labeled with ^{125}I using the glucose oxidase-lactoperoxidase method (Pink and Ziegler, (1979) Radiolabelling and characterization of cell surface molecules. In: Lefkovitz and Pernis (eds.), Research Methods in Immunology, pp. 169-180, NY Academic Press, Inc. Iodinations were performed in HBSS containing 20mM HEPES pH7.4. Cells were lysed at

$3x10^7$/ml for 30 minutes at 4°C immunoprecipitation buffer containing Ca^{++} ions (C-IPB) consisting of 1% Triton X-100, 50mM Tris pH7.4, 150mM NaCl, and 2mM $CaCl_2$. Leupeptin, antipapain, pepstatin, and chymostatin at 10µg/ml, soybean trypsin inhibitor at 20µg/ml and 1mM phenylmethylsulfonylfluoride were included as protease inhibitors. Lysates were centrifuged at 13,000xg for 15 minutes and precleared with Pansorbin cells (Behring Diagnostics, La Jolla, CA) or normal rabbit serum bound to protein A-Sepharose CL-4B. Rat monoclonal antibodies were bound to protein A-Sepharose CL-4B using a polyvalent rabbit antiserum to rat Ig (Pel Freez Biologicals, Rogers, AR). Immunosorbents were incubated with lysates for 3 hr at 4°C and washed in lysis buffer. Immunoprecipitates were analyzed by SDS-PAGE on 6% or 7% polyacrylamide gels. Molecular weight standards were myosin (M_r 200,000), ß-galactosidase (M_r116,000), phosphorylase b (M_r 97,000), bovine serum albumin (M_r 66,000), and ovalbumin (M_r 43,000). For some experiments monoclonal antibody R1-2 was purified on a goat anti-rat Ig column and covalently linked to Affigel 15 according to the manufacturers instructions (Bio-Rad Lab., Richmond, CA).

One-dimensional peptide mapping

Digestion of proteins with V8 protease form S.aureus (Sigma Chemical Co., St. Louis, MO) was carried out during gel electrophoresis (Cleveland et al., (1977) J. Biol. Chem. 252:1102-1106). After separation of proteins by SDS-PAGE, gel slices were excised and incubated for 15 min in 1mM EDTA, 0.1% SDS, 125mM Tris pH6.8. Gel slices were then loaded onto a 12.5% polyacrylamide gel and overlaid with 500 ng of V8 protease in 1mM EDTA, 0.1% SDS, 125mM Tris pH6.8, 20% glycerol containing bromphenol blue. Gel electrophoresis was interrupted for 1hr when the dye front

neared the end of the stacking gel to allow enzymatic digestion of proteins.

Isolation of RNA and Northern blot analysis

Cells were lysed in a 4M guanidinium isothiocyanate solution and the RNA was pelleted through a cushion of 5.7M CsCl (Chirgwin et al., (1979) Biochemistry 18:5294-5299). Poly(A^+) RNA was isolated by chromatography on oligo(dT) cellulose (type III, Collaborative Research). For each cell line, 4µg of denatured poly(A^+) RNA was separated on a 0.8% agarose/2.2M formaldehyde gel buffered with 40mM MOPS (pH7.5) and transferred to nylon membranes (Schleicher and Schuell). Probes were labeled to a specific activity of 2-4X10^8cpm/µg of DNA using the hexamer primer labeling procedure (Feinberg and Vogelstgein, (1983) Anal. Biochem. 132:6-13). The filters were hybridized at 42°C for 16hr in 3x SSPE, 50% formamide, 1x Denhardt's, 1% SDS and 100µg/ml herring testis DNA and washed in 0.2x SSPE, 0.1% SDS at 65°C. For hybridizations carried out under low stringency conditions, the formamide concentration in the hybridization solution was reduced to 35% and filters were washed in 2x SSPE, 0.1% SDS at room temperature. The cDNA clone pMINTß encoding amino acids 1-333 of murine integrin $ß_1$ was obtained from Drs. D.W. DeSimone, V. Patel, and R.O. Hynes. The cDNA clone pHFßA-1 containing human ß-actin (Gunning et al., (1983) Mol. Cell Biol. 3:787-795) was obtained from Drs. P. Gunning and L. Kedes, Stanford University Medical School. A 980bp ScaI/SalI fragment of pHFßA-1 was used for hybridizations.

Identification of a new integrin ß chain in a murine Peyer's patch homing receptor

The α subunit of LPAM-1 (hereafter called $α_{4m}$) has been shown to be analogous to the α chain of the

human integrin molecule VLA-4 as indicated below. The VLA-4 α chain is noncovalently associated with the integrin β_1 subunit. Whether the LPAM-1 β chain (hereafter called β_p) is analogous to β_1 was tested. Different rabbit antisera specific for β_1 did not recognize β_p or other proteins in lysates of surface labeled LPAM-1^+ TK1 lymphoma cells.

The analogy between the alpha subunit of LPAM-1 (hereafter called α_{4m}) and the alpha chain of the human integrin molecule VLA-4 was established as follows. A rabbit polyclonal antiserum specific for the alpha chain of human VLA-4 was tested for its ability to recognize the P160 subunit of LPAM-1. Immunoprecipitated SDS-denatured LPAM-1 was diluted in a buffer containing excess Triton X-100 and reanalyzed with different rabbit polyclonal antiserum. The monospecific rabbit anti VLA-4 alpha chain serum, but none of the control sera, specifically recognized P160 and its fragment P84. The rabbit anti VLA-4 serum was obtained by immunization with purified alpha chains and does not cross react with other integrin alpha subunits. Thus, on the basis of immunological cross reactivity, as well as structural similarities, the P160 subunit of LPAM-1 appears analogous, if not homologous, to the human VLA-4 alpha chain.

The anti VLA-4 antibodies immunoprecipitate a cell surface heterodimer of M_r 150,000 and 130,000, as well as two proteins of M_r 80,000 and 70,000, which were shown to be fragments of M_r 150,000 α chain protein. This is analogous to the alpha chain of the LPAM-1 antigen, which upon reducing conditions produces four proteins of apparent molecular weights of 160,000 (P160), 130,000 (P 130), 84,000 (P84), and 62,000 (P62). TK1 lymphoma cells, 3T3 fibroblasts or murine platelets were cell surface iodinated and immunoprecipitates were analyzed by SDS-page. The antibody used was R1-2. Immunoprecipitated material was treated

with 10 mM EDTA in 50 mM Tris pH7.4, 150 mM NaCl, 1% Triton X-100 and eluted material was analyzed using LPAM-1 heteroantiserum. All oth r immunoprecipitate were carried out from total cellular lysates: R1-2 LPAM-1 heteroantiserum, anti-LFA-1 alpha chain, anti-LFA-1 beta chain, anti-integrin β_1, anti-VLA-β, M1/70 normal rat serum, and normal rabbit serum. Samples were analyzed under reducing conditions. As a contro both rabbit antisera immunoprecipitated β_1 as well as an α chain of M_r 135,000 from murine 3T3 fibroblasts.

To characterize the β_p subunit, a polyvalent rat antiserum specific for LPAM-1 was obtained by immunization with immunoaffinity-isolated protein. As the assocation of LPAM-1 α and β subunits is dependent on the presence of Ca^{++} ions, β subunits can be selectively eluted with EDTA from LPAM-1 molecules bound to the α_{4m} specific antibody R1-2. The LPAM-1 heteroantiserum immunoprecipitated EDTA-eluted β_p subunits indicating that it contains antibodies recognizing β_p. Using this antiserum as well as antibody R1-2, LPAM-1 α or β chains were not detected on 3T3 fibroblasts. These results indicate that β_p is distinct from integrin β_1.

Next investigated was whether β_p is related to other integrin β chains. LFA-1 and LPAM-1 were both isolated from TK1 cells and their subunits compared by SDS-polyacrylamide gel electrophoresis. It was found that the β chain of LFA-1 (integrin β_2) and β_p could be clearly distinguished based on their molecular weights. Moreover, an antibody specific for β_2 did not cross-react with β_p or coprecipitate α_{4m} subunits. Conversely, the LPAM-1 heteroantiserum did not cross-react with or coprecipitate LFA-1 subunits. The β_p subunit was also compared to integrin β_3 which is identical to glycoprotein IIIa.

To compare the various β subunits in more detail, one-dimensional peptide mapping using V8

protease was carried out. The procedure was as follows. Gel slices containing radiolabeled beta subunits were excised after separation by SDS-page under non-reducing conditions and loaded onto a second polyacrylamide gel. Proteins were digested with 500ng V8 protease during the second electrophoresis. Proteins analyzed were LPAM-1 subunit β_p, integrin β_1, integrin β_2, and integrin β_3. The cellular sources of the various beta subunits were TK1 lymphoma cells, RAW112 lymphoma cells, 3T3 fiberblasts and murine platelets. The mapping showed that the digestion of β_p yielded peptide patterns clearly distinct from those of β_1, β_2, and β_3. Similarly, digestion of β_1, β_2, and β_3 each gave a unique peptide pattern. Therefore, these results support the concept that β_p represents a unique β subunit.

The LPAM-1 subunit β_p was further compared to β_1 by Northern blot analysis. Consistent with the absence of β_1 protein from TK1 cells a cDNA clone coding for an N-terminal fragment of murine β_1 did not hydridize with RNA from TK1 cells. The comparison was performed by isolating poly (A^+) RNA from β_p^+ β_1^- TK1 cells or $\beta_p^-\beta_1^+$ RAW 112 cells and hybridizing with cDNA clone pMINTβ encoding amino acids 1-333 of the murine integrin β_1 subunit or with a β-actin probe. Filters were hybridized and washed under low stringency or high stringency conditions. The hybridization with the beta-actin specific probe revealed that approximately equal amounts of TK1 and RAW112 poly (A^+) RNA were analyzed. The same results were obtained both under high and low stringency conditions. In contrast, the β_1-specific cDNA probe hydridized with two RNA species of the β_1^+ RAW112 lymphoma cells. These results clearly demonstrate that β_p is distinct from integrin β_1.

The VLA-4 like LPAM-1 α chain can associate with each of two different β chains

To investigate the nature of β subunits associated with α_{4m} on a panel of lymphoma cell lines, lysates of surface labeled lymphoma cells were immunoprecipitated using the α_{4m}-specific antibody R1-2 and analyzed by SDS-PAGE under non-reducing conditions. All cell lines expressed α chains of similar size as well as the two M_r 84,000 and 62,000 α chain fragments. However, a β_p-like subunit was only detected in cell lines TK1, TK23, and TK40. A protein of slightly higher apparent molecular weight (M_r 115,000 nonreduced) was coprecipitated from the cell lines apparently lacking β_p (i.e. TK50, L1-2, T69, KKT2). In addition, no β_p material was immunoprecipitated from most of these cell lines with the anti-LPAM-1 heteroantiserum. Both β_p and the M_r 115,000 protein were found to be coexpressed in cell lines TK23 and TK 40. The integrin β_1 subunit isolated with a monospecific rabbit antiserum from cell lines other than TK1 comigrated with the M_r 115,000 protein, but not with β_p. Immunoprecipitates using the β_1-specific antiserum also contained proteins comigrating with α_{4m}. An interpretation of these results is that α_{4m} can associate with either β_p or the M_r 115,000 protein, which may be identical to integrin β_1.

To test this hypothesis, the subunits associated with α_{4m}, were analyzed following immunoprecipitation from a panel of lymphoma cell lines using antibody R1-2 covalently linked to Affigel 15. A panel of lymphoma cell lines was cell surface iodinated and immunoprecipatated using the α_{4m} specific antibody R1-2 covalently linked to Affigel 15. Bound proteins were eluted with 100 ml of glycine pH 2.5, 1% Triton X-100 and eluates were diluted 1:5 with 50 mM Tris pH 8.8, 150 mM NaCl, 10 mM EDTA, 1% Triton X-100. Eluted and associated subuits were split in several aliquots and

reanalyzed with anti-integrin β_1, anti-VLA-β, the LPAM-1 heteroantiserum, and monoclonal antibody 30G12 directed against the leucocyte common antigen T200. Samples were analyzed by SDS-PAGE under non-reducing conditions. The bound proteins were eluted, and the dissociated subunits were reanalyzed using the LPAM-1 heteroantiserum or two antisera directed against β_1. Consistent with the results presented above, the β_p subunit expressed in TK1 cells reacted with the anti-LPAM-1 heteroantiserum, but not with the β_1-specific antisera. In contrast, the M_r 115,000 protein coprecipitated by antibody R1-2 from β_p-negative cell lines TK50, L1-2, T69, and KKT2 was recognized by both β_1-specific antisera, but not by the LPAM-1 heteroantiserum. Both β_p and the M_r 115,000 protein were isolated from cell lines TK23 and TK40. Accordingly, both anti-β_1 antisera specifically reacted with the M_r 115,000 protein, whereas β_p was only detected by the LPAM-1 heteroantiserum. These results show directly that β_p as well as a distinct M_r 115,000 protein (integrin β_1) are copurified with an antibody specific for α_{4m}.

Gel slices containing radialabeled α_{4m} were excised from an SDS-PAGE (nonreducing) separation and digested with 500 ng V8 protease during the electrophoreses in a second polyacrylamide gel. The α_{4m} subunits were isolated from cell lines TK1, TK23, TK40, and TK50. The identity of the α subunits recognized by antibody R1-2 was further verified by one-dimensional peptide mapping. Digestion of α chains isolated from four different cell lines with V8 protease yielded identical peptide patterns regardless of their association with β_p or β_1 indicating that antibody R1-2 recognized the same α on different cell lines. Therefore the VLA-4-like LPAM-1 α chain is the common subunit of two distinct cell surface heterodimers: LPAM-1, composed of α_{4m} associated with

β_p, and LPAM-2, consisting of α_{4m} and integrin β_1.

Both LPAM-1 and LPAM-2 are involved in lymphocyte-Peyer's patch HEV interactions

The cellular distribution and function of both LPAM-1 and LPAM-2 heterodimers were investigated. The presence of LPAM-1 and LPAM-2 was determined by immunoprecipitation with the α_{4m}-specific antibody R1-2 and subsequent analysis of the β subunits with the LPAM-1 heteroantiserum or β_1 specific antisera (described above). The binding capacity of cells for HEV in Peyer's patches or peripheral lymph nodes was tested in a modified Stamper & Woodruff in vitro assay. Results showed that all Peyer's patch HEV-binding cell lines as well as a subset of non-binding lymphomas reacted with antibody R1-2. When analyzed by immunoprecipitation, the Peyer's patch HEV-binding lymphomas showed a heterogeneous expression of LPAM-1 and LPAM-2. Whereas both heterodimers were coexpressed in cell lines TK23 and TK40, other cell lines were found to be singly positive for either LPAM-1 (cell line TK1) or LPAM-2 (cell line TK50). In normal mesenteric node lymphocytes both heterodimers were detected. In contrast to Peyer's patch HEV-binding cell lines, all R1-2-reactive non-binding lymphomas expressed LPAM-2, but not LPAM-1. Antibody R1-2 inhibited the binding of all lymphoma cell lines tested to Peyer's patch HEV but not to peripheral lymph node HEV consistent with previous results indicating that it recognized a murine Peyer's patch homing receptor. As antibody R1-2 also blocked the adhesion of the LPAM-1 or LPAM-2 single-positive lymphomas TK1 and TK50 Peyer's patch HEV, these results further suggest that, in addition to LPAM-1, LPAM-2 is also involved in lymphocyte-Peyer's patch HEV interactions.

The above results demonstrate that novel proteins may be employed for specific binding to

particular anatomical sites. The different proteins may be used in a variety of ways to prevent cells from binding or to direct compositions to the desired sites. In this manner, the immune system may be modulated by increasing or decreasing lymphocyte populations at specific sites. The ability to control the lymphocyte population at particular sites, may be used to protect against autoimmune diseases, reduce the inflammatory response, to localize specific cells or drugs for diagnosis or therapy for neoplastic conditions, and to enhance immune responses by modifying viruses which may be endocytosed by lymphocytes or monocytes for presentation to T-cells.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

WHAT IS CLAIMED IS:

1. A DNA sequence encoding a homing receptor unit selected from the group consisting of α_{4m}, b_p or the core protein $gp90^{Mel-14}$ free of ubiquitin or an individual domain thereof, other than as part of a mammalian chromosome.

2. A DNA sequence according to Claim 1, wherein said DNA is cDNA.

3. A DNA sequence according to Claim 2, wherein said unit is a_{4m} or b_p.

4. A DNA sequence according to Claim 2, wherein said unit is the core protein $gp90^{Mel-14}$ or a domain comprising the signal sequence, lectin-like domain, EGF-like domain or the complement regulatory protein-like domain.

5. A DNA sequence comprising a DNA sequence encoding a homing receptor unit selected from the group consisting of β_p or $gp90^{Mel-14}$ or domain thereof, other than as part of a mammalian chromosome joined to at least one of other than the wild-type transcriptional initiation region, a marker or a replication system for stable replication in a cellular host.

6. A method for modulating homing of a component of interest to a homing ligand of a high endothelial venule associated with a mucosal membrane lymphoid organ or tissue of a mammalian host, providing for binding to or inhibiting binding to said ligand, said method comprising:

administering to said host a homing modulating amount of a composition comprising an

antibody to LPAM-1, -2 or VLA-4, a peptide of LPAM-1, -2, VLA-4, or core protein of $gp90^{Mel-14}$ or domain thereof, capable of binding to said mucosal membrane lymphoid organ or tissue ligand or lymph node, a peptide immunologically cross-reactive therewith, or a conjugate thereof, to modulate binding to mucosal membrane lymphoid organ or tissue ligand;

whereby said composition modulates the binding to said ligand.

7. A method according to Claim 6, wherein said composition comprises an antibody to LPAM-1, -2 or VLA-4, a peptide of LPAM-1, -2, VLA-4, core protein of $gp90^{Mel-14}$ or extracytoplasmic domain thereof, capable of binding to said mucosal membrane lymphoid organ or tissue ligand.

8. A method according to Claim 7, wherein said composition comprises VLA-4 or fragment thereof capable of binding to said mucosal membrane lymphoid organ or tissue ligand.

9. A method according to Claim 8, wherein said composition comprises LPAM-1 or -2.

10. A composition comprising at least about 50 wt.% of LPAM-1 and/or -2.

11. A composition comprising at least about 50 wt.% of a mammalian α_{4m} or b_p.

12. A composition comprising a fragment of LPAM-1, or -2 capable of binding to a mucosal membrane high endothelial venule.

13. A composition comprising a fragment of a mammalian α_{4m} or b_p of at least about 8 amino acids.

14. A DNA sequence of at least about 12nt having at least about a 95% identity with a sequence of the gene encoding α_{4m} or b_p and terminating at the coding sequence or joined to other than the natural contiguous DNA.

15. A DNA sequence according to Claim 14, comprising a cDNA.

16. A DNA sequence encoding α_{4m}.

17. A DNA sequence encoding b_p.

18. A DNA sequence of at least about 12nt having at least about a 95% identity with a sequence of the gene encoding $gp90^{Mel-14}$ and terminating at the coding sequence or joined to other than natural contiguous DNA.

19. A DNA sequence according to Claim 18, comprising a cDNA sequence.

20. A DNA sequence according to Claim 18, comprising the signal sequence, lectin-like domain, EGF-like domain or the complement regulatory protein domain.

21. Antibodies to the core protein of $gp90^{Mel-14}$ capable of blocking binding to high endothelial venules.

22. Antibodies to a_{4m} or b_p capable of blocking binding to high endothelial venules.

23. A cell comprising a construct comprising a DNA sequence according to any of Claims 14 or 18 under the transcriptional and translational regulation

of regulatory regions functional in said cell, wherein said construct is present in said cell as a result of introduction of said construct into said cell.

24. A method for inhibiting metastasis to a high endothelial venule site, said method comprising:

administering to a mammalian host a binding inhibiting amount of a composition comprising an antibody to LPAM-1, -2 or VLA-4, a peptide of LPAM-1, -2, VLA-4, or core protein of $gp90^{Mel-14}$ capable of binding to said mucosal membrane lymphoid organ or tissue ligand or lymph node, a peptide immunologically cross-reactive therewith, or a conjugate thereof.

25. A method for directing a cell or virus to a mucosal membrane lymphoid organ or tissue or lymph node in a mammalian host, said method comprising:

introducing into said cell or genome of said virus an expression cassette comprising a DNA sequence according to Claim 5 or functional fragment thereof for expression of said unit or functional fragment thereof, whereby the expression product occurs on the surface of said cell or virus and is able to bind to a high endothelial venule ligand, to produce cells or viruses which home to said high endothelial venules.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/00, C07 K 15/00
A61K 39/00

A3

(11) International Publication Number: WO 90/07321

(43) International Publication Date: 12 July 1990 (12.07.90)

(21) International Application Number: PCT/US89/05067

(22) International Filing Date: 10 November 1989 (10.11.89)

(30) Priority data:

289.201	23 December 1988 (23.12.88)	US
315.736	24 February 1989 (24.02.89)	US

(71) Applicant: THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY [US/US]: Stanford University, Stanford, CA 94305 (US).

(72) Inventors: WEISSMAN, Irving, L. ; 1170 Welch Road, Apt. 711, Stanford, CA 94305 (US). HOLZMANN, Bernard ; Gufidaunerstr. 30, D-8000 Münich 90 (DE). SIEGELMAN, Mark, H. ; 2035 Columbia Avenue, Palo Alto, CA 94306 (US).

(74) Agent: LEYDIG, VOIT & MAYER; 350 Cambridge, Suite 200, Palo Alto, CA 94306 (US).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
9 August 1990 (09.08.90)

(54) Title: LYMPHOCYTE RECEPTOR HOMING SEQUENCES AND THEIR USES

(57) Abstract

Proteins are identified as homing receptors for Peyer's patches and lymph nodes, where the proteins may be used for inhibiting homing of lymphocytes or providing for homing of drugs or other compositions for *in vivo* diagnosis or therapy. In addition, nucleic acid compositions are provided which may be used for expression of the proteins or fragments thereof or for transforming cells to provide for enhanced homing capability or for inhibiting or modulating such homing.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CM	Cameroon
DE	Germany, Federal Republic of
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
HU	Hungary
IT	Italy
JP	Japan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
MC	Monaco
MG	Madagascar
ML	Mali
MR	Mauritania
MW	Malawi
NL	Netherlands
NO	Norway
RO	Romania
SD	Sudan
SE	Sweden
SN	Senegal
SU	Soviet Union
TD	Chad
TG	Togo
US	United States of America

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US89/05067

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC (5) C12N 15/00; C07K 15/00; A61K 39/00
U.S. CL. 536/27; 530/328-324; 424/85.8,88

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
U.S. CL.	530/387 536/27 435/172.3,69.1,235,320 435/240.1, 530/328-324,350 424/85.8,88

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [10]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
Y,P	Cell, Volume 56, Issued 24 March 1989, (Johnston et al.) "Cloning of GMP-140, a Granule Membrane Protein of Platelets and Endothelium: Sequence Similarity to Proteins Involved in Cell Adhesion and Inflammation" See pages 1033-1044, particularly figures 1-9 and Table I.	1-13,16-17 21-22 and 24-25
X,P Y,P	Cell, Volume 56, Issued 24 March 1989, (Goldstein et al.) "A Human Lymphocyte Homing Receptor, the Hermes Antigen is Related to Cartilage Proteoglycan Core and Link Proteins" See pages 1063-1072, particularly figures 4-7.	1-5,10-13,16-17 6-9,21-22, 24-25
X,P Y,P	Cell, Volume 56, issued 24 March 1989, (Stamenkovic et al.) "A Lymphocyte Molecule Implicated in Lymph Node Homing Is a Member of the Cartilage Link Protein Family" See pages 1057-1062, particularly figure 4.	1-5,10-13,16-17 6-9,21-22 24-25
X,P Y,P	Cell, Volume 56, Issued 24 March 1989, (Lasky et al.) "Cloning of a Lymphocyte Homing Receptor Reveals a Lectin Domain" See pages 1045-1055, particularly figures 2,4, and 5.	1-5,10-13,16-17 6-9,21-22 24-25

(continue on extra sheets)

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search

10 MAY 1990

Date of Mailing of this International Search Report

12 JUN 1990

International Searching Authority

ISA/US

Signature of Authorized Officer

MICHELLE S. MARKS-PATENT EXAMINER

Form PCT/ISA/210 (second sheet) (Rev.11-87)

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No. [18]
Y	Cell, Volume 47, Issued 05 December 1986, (Kikutani et al.) "Molecular Structure of Human Lymphocyte Receptor for Immunoglobulin E" See pages 657-665, particularly figure 5.	6-9 and 24
X,P Y,P	EMBO Journal, Volume 8 No.5, Issued May 1989, (Takada et al.)"The primary structure of the a4 subunit of VLA-4:homology to other integrins and a possicle cell-cell adhesion function" See pages 1361-1368, particularly figures 1-7 and Tables 1-2.	1-5,10-13,16-17 6-9,21-22, 24-25
X Y	Gene, Volume 52 No.1, Issued 1987, (Prat et al.) "Multiple variability in the sequence of a family of maize endosperm proteins" See pages 41-49 particularly the peptide sequence PQCQ in figure 3.	10-13 6-9 and 24
X Y	The Journal of biological Chemistry, Volume 262,No.22, Issued 05 August 1987, (Pikkarainen et al.) "Human Laminin B1 Chain" See pages 10454-10462 particularly the peptide AGYYG in figures 2 and 5.	10-13 6-9 and 24
X Y	The Journal of Biological Chemistry, Volume 263,No.13, Issued 05 May 1988, (Nakanishi et al.) "Complete Nucleotide Sequence and Characterization of the 5'-Flanking Region of Mammalian Elongation Factor 2 Gene" See pages 6384-6391 particularly the nucleotide sequence ATGCTCTGT and TTGTGGGG in figure 3.	1-5, 16-17 6-9,21-22,24 10-13 and 25
X Y	The Journal of Biological Chemistry, Volume 261,No.15, Issued 25 May 1986, (Drickamer et al.)"Mannose-binding Proteins Isolated from Rat Liver Contain Carbohydrate-recognition Domains Linked to Collagenous Tails" See pages 6878-6887 particularly the peptide GEPNN and EDCV in figures 10, 7 and 8.	10-13 6-9 and 24
X Y	The Journal of Biological Chemistry, Volume 261,No.18, Issued 25 June 1986, (Doege et al.)"Partial cDNA Sequence Encoding a Globular Domain at the C Terminus of the Rat Cartilage proteoglycan" See pages 8108-8111.	10-13 6-9 and 24
X Y	The Journal of Biological Chemistry, Volume 262,No.36, Issued 25 December 1987, (Doege et al.)"Complete Primary Structure of the Rat cartilage Proteoglycan Core Protein Deduced from cDNA Clones. See pages 17757-17767, particularly the peptide ETSAS in figure 3.	10-13 6-9 and 24
X Y	The Journal of Immunology, Volume 138,No.12, Issued 15 June 1987, (Lewinsohn et al.)"Leukocyte-Endothelial Cell Recognition: Evidence of a Common Molecular Mechanism Shared By Neutrophils, Lymphocytes, and Other Leukocytes" See pages 4313-4321, particularly Table 1 and 2 and figure 7.	21,22 6-9 and 24

(continued on next page)

III. DOCUMENTS CONSIDERED TO BE RELEVANT (NTINUED FROM THE SEC ND SHEET)

Category *	Citation of Document, 16 with indication, where appropriate, of the relevant passages 17	Relevant to Claim No 18
X,P Y,P	The Journal of Immunology, Volume 143,No.10, Issued 15 November 1989 (Jutila et al.)"Function And Regulation Of The Neutrophil MEL-14 Antigen In Vivo: Comparison With LFA-1 and MAC-1" See pages 3318-3324, particularly Tables V-VII.	10-13,21-22 6-9 and 24
X,P	The Journal of Immunology, Volume 143,No.10, Issued 15 November 1989, (Zhou et al.)"Molecular Cloning and expression of Pgp-1" See pages 3390-3395, particularly figures 1 and 2.	10-13
X Y	Journal of molecular Biology, Volume 196, Issued 12 January 1987, (Van het Schip et al.)"Nucleotide Sequence of the Encoded Yolk Precursor Protein" See pages 245-260 particularly the peptide SVLS in figures 2 and 5.	10-13 6-9 and 24
X Y	Nature, Volume 304, Issued 07 July 1983, (Gallatin et al.)"A cell-surface molecule involved in organ-specific homing of Lymphocytes" See pages 30-34, particularly Table 1 and Figure 4 showing 90 KD molecular size of gp90MEL-14.	6-9 and 24 21,22
X Y	Nature, Volume 309, Issued 31 May 1984, (Ullrich et al.)"Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells" See pages 418-425 particularly the peptide WGAGGE in figure 2.	10-13 6-9 and 24
Y	Nature, Volume 309, Issued 28 June 1984, (Xu et al.) "Human epidermal growth factor receptor cDNA is homologous to a variety of RNAs overproduced in A431 carcinoma cells" See pages 806-810.	6-9 and 24
Y	Nature, Volume 309, Issued 17 May 1984,(Mroczkowski et al.)"ATP-stimulated interaction between epidermal growth factor receptor and supercoiled DNA" See pages 270-273 particularly the peptide WGAGE.	6-9 and 24
X	Nature, Volume 320, Issued 13 March 1986, (Noda et al.)"Existence of distinct sodium channel messenger RNAs in rat brain" See pages 188-192.	10-13
X	Nucleic Acids Research, Volume 12,No.2, Issued 1984, (Burch et al.)"Identification and sequence analysis of the 5' end of the major chicken vitellogenin gene" See pages 1117-1135.	10-13
X Y	Proceedings of the National Academy of Sciences USA, Volume 78,No.2, Issued February 1981, (Shinoda et al.) "Complete amino acid sequence of the Fc region of a human o chain" See pages 785-789 particularly the peptide TQPL in figures 1 and 2.	10-13 6-9 and 24

(continue on next page)

III. DOCUMENTS CONSIDERED T BE RELEVANT (C NTINUED FROM THE SEC ND SHEET)

Category *	Citation of Document, 16 with indication, where appropriate, of the relevant passages 17	Relevant to Claim No 1
X	Science, Volume 224, Issued 25 May 1984, (Lin et al.) "Expression Cloning of Human EGF Receptor Complementary DNA: Gene Amplification and three Related Messenger RNA Products in A431 Cells" See pages 843-848.	10-13
X,P	Science, Volume 243, Issued 03 March 1989, (Bevilacqua et al.)"Endothelial leukocyte Adhesion Molecule 1_ An Inducible Receptor for Neutrophils Related to Complement regulatory Proteins and Lectins" See pages 1160-1165, particularly figures 1-6.	1-5,16-17 21,22 and 25
X Y	Science, Volume 231, Issued 21 February 1986, (Siegelman et al.)"Cell Surface Molecule Associated with Lymphocyte Homing Is Ubiquitinated Branched-Chain Glycoprotein" See pages 823-829, particularly figures 3,5,6,8 and 10.	10-13 1-5,16-17, 6-9,24-25
X,P Y,P	Science, Volume 243, Issued 03 March 1989, (Siegelman et al.)"Mouse Lymph Node Homing Receptor cDNA Clone Encodes a Glycoprotein Revealing Tandem Interaction Domains" See pages 1165-1172, particularly figures 1-10.	1-5,6-9 16-17,24 10-13
X Y	Science, Volume 231, Issued 21 February 1986, (St. John et al.)"Expression Cloning of a Lymphocyte Homing Receptor cDNA: Ubiquitin Is the Reactive Species" See pages 845-850 particularly figure 1 and Table 1.	1-5,6-9 16-17,24 25
X Y	Ubiquitin, edited by Martin Rechsteiner, Issued 1988, Plenum Publishing Corporation (Siegelman et al.) "Lymphocyte Homing Receptors, Ubiquitin, and Cell Surface Proteins" See pages 239-269, particularly figures 5-7.	10-13 1-5,6-9 16-17 and 24-25

FURTHER INFORMATI N CONTINUED FR M THE SECOND SHEET

V. ☒ OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers . because they relate to subject matter [2] not required to be searched by this Authority, namely:

2. ☒ Claim numbers 14-15,18-20 and 23 . because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out [3], specifically:

See continuation sheets

3. ☐ Claim numbers_____, because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. ☒ OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this international application as follows:

See continuation sheets

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:
1-13,16-17,21-22 and 24-25 to the extent they read on the species for which fees were paid

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

☒ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional search fees.

V. OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE

Claims 14-15, 18-20 and 23 fail to comply with the prescribed requirements to such an extent that a meaningful search could not be carried out. See PCT Article 17 (2)(a)(ii). Applicants comments on page 2 of their protest of 24 April 1990 states that these claims are directed to specific sequence of the DNA sequences of claim 1. This statement is incorrect, as no specific sequence is elucidated by a statement of "A DNA sequence of at least about 12nt [nucleotides] having at least about a 95% identity with a sequence of the gene encoding 4m or bp and terminating at the coding sequence or joined to other than the natural contiguous DNA.". Please note that the DNA sequence of the LPAM-1 alpha (i.e. 4m) and beta (i.e. p) subunits is never defined in the instant specification. A meaningful search cannot be carried out on such an open, vague and non-defined claim.

VI. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

The inventions are grouped above according to the unity of invention concept reflected in Rule 13.2.

I. Claims 1-5, drawn to a DNA sequence encoding the 4m homing receptor unit, classified in Class 536, subclass 27.

Claims 1-5 are generic to a plurality of disclosed patentably distinct species comprising DNA encoding:a)gp90mel14 free of ubiquitin, b)bp, c) gp90mel-14 core, d) gp90mel14signal sequence, e) gp90mel14 lectin-like domain, f) gp9014core complement regulated protein likme domain or g) a specific 4m oligonucleotide or h) a specific bp oligonucleotide.

UNITY

II. Claims 6-9, drawn to a method of modulating homing of a component of interest to a homing ligand of a high endothelial venule associated with a mucosal membrane, classified in Class 436, subclass 501. Claims 6-9 are generic to a plurality of disclosed patentably distinct species comprising the above method employing the use of :
i)antibody to LPAM-1, j) antibody to LPAM-2, k) antibody to VLA-4, l) a specific LPAM-1 peptide m) a specific LPAM-2 peptide n) a specific VLa-4 peptide, o) a specific gp90mel'4 core protein, p) a specific gp90mel14 domain q) the extracytoplasmic domain of gpmel14 r) a specific VLA-4 fragment.

III. Claims 10-13, drawn to a composition comprising either LPAM-1, LPAM-2, 4m, bp, LPAM-1 fragment, LPAM-2 fragment , 4m fragment of at least 8 amino acids, or bp fragment of at least 8 amino acids, classified in Class 530, subclasses324-328 and 350 and 387.

Claims 10-13 are generic to a plurality of disclosed patentably distinct species comprising at lest about 50 wt.% of the following :
s) LPAM-1, t) LPAM-2 u) 4m v) bp w) a specific LPAM-1 fragment x) a specific LPAM-2 ffragment y) a specific 4m fragment of at least 8 amino acids z) a specific bp fragment of at least 8 amino acids.

IV. Claims 21-22, drawn to an antibody, classified in class 530 subclass 387.

Claims 21-22 are generic to a plurality of disclosed patentably distinct species comprising : aa)antibody specific for gp90mel14 core protein, bb) antibody specific for 4m and cc) antibody specific for bp.

UNITY

V. Claim 24, draw to a method for inhibiting metastasis to a high endothelial venule site using either antibodies or peptides classified in class 424 subclass 85.

Claim 24 is generic to a plurality of disclosed patentably distinct species comprising the above method employing the use of : dd)antibody to LPAM-1, ee) antibody to LPAM-2, ff) antibody to VLA-4, gg) a specific LPAM-1 peptide hh) a specific LPAM-2 peptide ii) a specific VLa-4 peptide, JJ) a specific gp90mel14 core protein, kk) a specific gp90mel14 domain ll) the extracytoplasmic domain of gpmel14 mm) a specific VLA-4 fragment or nn) a specific immunologically cross-reactive peptide or oo) a specific conjugate employing one of dd-oo above.

VI. Claim 25, drawn to a method for directing a cell or virus to a mucosal membrane lymphoid organ or tissue or lymph node in a mammalian host using a DNA expression system, classified in class 424 88 and 89.

Claim 26 generic to a plurality of disclosed patentably distinct species comprising the above method employing the use of DNA encoding the following : pp)gp90mel14 free of ubiquitin, qq)bp, rr) gp90mel-14 core, ss) gp90mel14signal sequence, tt) gp90mel14 lectin-like domain, uu) gp9014core complement regulated protein likme domain or vv) a specific 4m oligonucleotide or ww) a specific bp oligonucleotide.